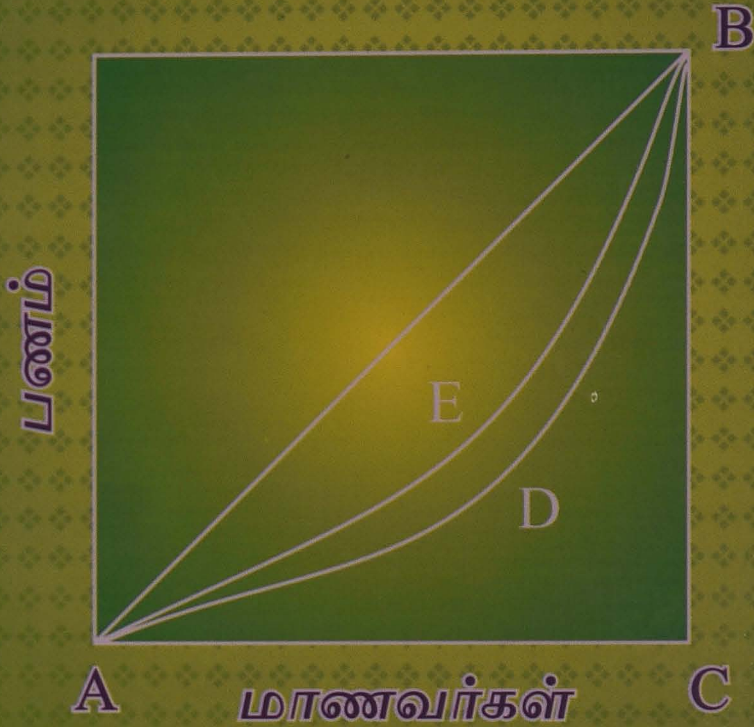
புள்ளியியல் முறைகள்



முனைவர் ச. அய்யம்பிள்ளை



தமிழ்நாடு மாநில உயர்கல்வி மன்றம்

சென்னை - 600 005.

புள்ளியியல் முறைகள்

முனைவர். ச. அய்யம்பிள்ளை



தமிழ்நாடு மாநில உயர்கல்வி மன்றம்
சென்னை - 600 005.

முதற்பதிப்பு : 2011

பதிப்புரிமை : தமிழ்நாடு மாநில உயல்கல்வி மன்றம்
சென்னை - 600 005

நூலின் பெயர் : புள்ளியியல் முறைகள்

நூலாசிரியர் : முனைவர் ச. அய்யம்பிள்ளை,
பேராசிரியர்,
பொருளியல் துறை,
பாரதிதாசன் பல்கலைக்கழகம்,
காசாமலை வளாகம்,
திருச்சிராப்பள்ளி - 620 023.

மறு ஆய்வு செய்தவர்கள் : முனைவர் ஆர். இளங்கோ,
பொருளியியல் துறை தலைவர் (ஓய்வு),
அண்ணாமலைப் பல்கலைக்கழகம்,
அண்ணாமலைநகர் - 608 002.



: முனைவர் டி. கோவிந்தராஜ்,
பேராசிரியர் மற்றும் தலைவர்,
பொருளாதாரத்துறை,
ஸ்ரீ புஷ்பம் கல்லூரி,
பூண்டி, தஞ்சாவூர் மாவட்டம்.

தமிழ் திருத்தம் செய்தவர் : முனைவர் தங்கமாரிமுத்து,
துறைத் தலைவர்,
தமிழ்த்துறை,
S.I.V.E.T. கல்லூரி
கவுரிவாக்கம்,
சென்னை - 600 073.

விலை : ரூ. 100.00

அச்சிட்டோர் : பவர்மேன் பிரிண்டர்ஸ்
எண்.6/15, டாக்டர் ராதாகிருஷ்ணன் நகர்,
3வது தெரு, கொருக்குப்பேட்டை
சென்னை - 600 021.
செல் : 98846 99888

புள்ளியியல் முறைகள்
(STATISTICAL METHODS)

பொருளடக்கம்

முன்னுரை

புள்ளியியல் முறைகள் பற்றிய பல புத்தகங்களைப் படிக்கும்பொழுது பலவிதமான தொடரமைப்புகள் காணப்படுகின்றன. இவற்றில் எந்தவிதமான அமைப்பு எங்கு பொருத்தமானதாக இருக்கும் என்பதில் பல கருத்து வேறுபாடுகள் காணப்படுகின்றன. அவற்றை மிகவும் தோராயமாகப் பார்த்தால், இரண்டுவித தொடரமைப்புகள் (Arrangements) தெளிவாகத் தெரிகின்றன. ஒன்று புள்ளியியல் முறைகள் பாடமாகப் போதிக்கப்படவேண்டும் என்ற எண்ணத்துடன் எழுதப்பட்ட புத்தகங்கள். மற்றொன்று புள்ளியியுல் முறைகள் ஆய்வுக் கருவிகளாகப் போதிக்கப்பட வேண்டும் என்ற எண்ணத்தில் எழுதப்பட்ட புத்தகங்கள். இந்தப் புத்தகம் இரண்டாம் வகையை ஒட்டியிருப்பதற்காக தொடர்ந்தமைக்கப்பட்டுள்ளது. எனவே ஆய்வாளரின் தேவைக்கேற்பப் பாடங்கள் வரிசைப்படுத்தப்பட்டுள்ளன. புள்ளிவிபரங்களின் தன்மைகளும், அவற்றை சேகரிக்கும் முறைகளும் இப்புத்தகத்தின் முதற்பாடத்திலேயே விளக்கப்படுகின்றன. சில புத்தகங்களில் வேறு மாதிரியான வரிசைப்படுத்துதலும் காணப்படலாம்.

புள்ளியியல் முறைகளில் பொதிந்திருக்கும் பல கருத்துருக்களின் ஆங்கிலப் பதங்களுக்கு பொருத்தமான தமிழ்ப் பதங்கள் கண்டுபிடிப்பதில் சில சிக்கல்கள் உணரப்பட்டன. அகராதிகளிலும் தமிழிலான புத்தகங்களிலும் பொருத்தமான, சமமான தமிழ்ப் பதங்கள் தேடியபோது ஒரு ஆங்கிலப் பதத்திற்கு ஒன்றிற்கு மேற்பட்ட தமிழ்ப் பதங்களும், ஒரு தமிழ்ப் பதத்திற்கு ஒன்றிற்கு மேற்பட்ட ஆங்கிலப் பதங்களும் சரியாக இருப்பதுபோல் தோன்றுகின்றன. ஆனாலும், பல சமயங்களில் பல

பதங்களுக்கு அந்தப் பதங்களைப் பயன்படுத்தப்பட வேண்டிய இடங்களுக்குப் பொருத்தமான பதங்கள் கிடைக்கவில்லை என்பதே உண்மை. எனவே பொருத்தமான பொருள்களைக் கொண்டு பதங்கள் இப்புத்தகத்தில் கையாளப்பட்டுள்ளன. வேறு புத்தகங்களில் வேறு பதங்களும் காணப்படலாம். இப்புத்தகத்தில் விபரங்கள் என்னும் சொல் பல இடங்களில் விவரங்கள் என்றும் பயன்படுத்தப்பட்டுள்ளது.

புள்ளியியலில் பல மொழிகளில் இருந்து எடுக்கப்பட்டுள்ள குறியீடுகள் (notations and symbols) பயன்படுத்தப்பட்டுள்ளன. வழக்கத்தில் உள்ள குறியீடுகளைப் பற்றி அறிந்தவர்களுக்கு, அவை குறிக்கும் கருத்தினைப் புரிவது சுலபமாக இருக்கும். குறியீடுகள் எவ்வாறு புரியப்பட்டனவோ, அதேபோல் புரிய முயற்சி செய்யும்போது மூளைக்கு ஏற்படும் அயற்சியையும் சோர்வினையும் குறைத்து புரிதலை எளிதாக்கலாம். எனவே, வழக்கத்தில் உள்ள குறியீடுகளே இங்கும் பயன்படுத்தப்பட்டுள்ளன.

மிகவும் பரவலாகப் பேசப்பட்டு வரும் கருத்துகள் இந்தப் புத்தகத்தில் சுருக்கமாக விளக்கப்பட்டுள்ளன. அதேசமயம், தவறாகப் புரிந்து கொள்ள வாய்ப்பு அதிகமாக இருக்கின்ற கருத்துகள் சற்று அதிக விளக்கத்துடன் அளிக்கப்பட்டுள்ளன. புள்ளியியல் பட்டியல்கள் (Statistical Tables) வெவ்வேறு புத்தகங்களில் வெவ்வேறுவிதமாகத் தரப்பட்டுள்ளன. அதனால், அவற்றைப் புரிவதில் குழப்பங்கள் விளைகின்றன. அதைத் தவிர்ப்பதற்காக, ஒவ்வொரு பட்டியலும் வெவ்வேறு விதமாக இருந்தால், அவற்றை அருகருகே இந்தப் புத்தகத்தில் வைத்து, அவற்றிற்கிடையே உள்ள வேறுபாடுகளைப் புரிந்து கொள்ள வழிவகை செய்யப்பட்டுள்ளது.

ச.அய்யம்பிள்ளை

1. புள்ளியியல் - ஓர் அறிமுகம்

புள்ளி விபரங்களை (Statistics) எப்படிச் சேகரிப்பது, ஒழுங்குபடுத்துவது, வகைப்படுத்துவது, அட்டவணைப் படுத்துவது, வரைபடங்களாக்குவது, படிப்பவர்களுக்கு எளிதாக விளங்க வைப்பது, ஆய்வு செய்வது, பொருள் அறிவது, விவரிப்பது போன்ற செய்திகளை உள்ளடக்கியது தான் புள்ளியியல் (Statistics) முறைகள். எல்லாவகையான ஆய்வுகளும் பலவகையான புள்ளி விபரங்களை வைத்தேதான் நடக்க முடியும்; நடக்கின்றன. எனவே பாட பேதமின்றி ஆய்வுகள் மேற்கொள்ளும் அனைவரும் புள்ளியியல் முறைகளைக் கற்றுக் கொள்ள வேண்டிய கட்டாயத்தில் இருக்கின்றார்கள். மேலும், இன்றைய காலகட்டங்களில் கணினிகள் வாழ்வின் போக்கினை நிர்ணயிக்கின்ற அளவுக்கு வியாபித்துள்ளதால் புள்ளி விபரங்களை எளிதாக ஆய்வு செய்வதற்கும் கணினிகள் பயன்படுத்தப்படுகின்றன. எனவே, எண்ணாலல்லாத விபரங்களும் (எ.கா. சாதி, மதம், நேர்மை, குணம்) இப்போது எண்களாக்கப்பட்டு (Proxy, Dummy) கணினிக்குள் புகுத்தப்படுகின்றன. இவ்வாறாகப் புள்ளியியல் முறைகளின் முக்கியத்துவம் கூடிக்கொண்டே வருகிறது.

புள்ளி விபரங்கள்

புள்ளி விபரங்கள் அவற்றின் தன்மைகளைப் பொறுத்து பலவிதமாக வகைப்படுத்தப்படுகின்றன.

முதலில் யாரால் புள்ளி விபரங்கள் சேகரிக்கப் படுகின்றன என்பதைப் பொறுத்து முதல்நிலை மற்றும் இரண்டாம் நிலைப் புள்ளி விபரங்கள் என பிரிக்கப்படுகின்றன. ஆய்வாளர்கள் தங்கள் தேவைக்கேற்ப தாங்களே படிக்கப்பட இருக்கின்றவர்களிடம் அவர்களைப் பற்றிய விபரங்களைச் சேகரித்தால் அவை முதல்நிலைப்

புள்ளி விபரங்கள் என அழைக்கப்படுகின்றன. அப்படியின்றி முன்னரே வேறு சிலரால் வேறு சில காரணங்களுக்காகச் சேகரிக்கப்பட்ட விபரங்களை பிறகு சிலர் பயன்படுத்தினால் அப்படிப்பட்ட புள்ளி விபரங்கள் இரண்டாம் நிலைப் புள்ளி விபரங்கள் என அழைக்கப்படுகின்றன. முதல் நிலைப் புள்ளி விபரம் என்றால் முதலில் சேகரிப்பது என்றும் இரண்டாம் நிலைப் புள்ளி விபரம் என்றால் இரண்டாவதாகச் சேகரிப்பது என்றும் பொருள் கொண்டுவிடக்கூடாது. உண்மையாகச் சொல்லப் போனால் ஆய்வுப் பொருள் தொடர்பான இரண்டாம்நிலைப் புள்ளிவிபரங்களை நன்கு படித்த பின்னர் தான் முதல்நிலைப் புள்ளிவிபரங்களைச் சேகரிக்கச் செல்ல வேண்டும்; அப்படியில்லையானால் முதல்நிலைப் புள்ளி விபரம் முழுமை பெற்றதாகவும் இருக்காது; பயனுள்ளதாகவும் இருக்காது.

எண்களால் தரப்படக்கூடிய புள்ளி விபரங்கள் என்றும் எண்களால் தரப்பட முடியாத குணம் தழுவிய புள்ளி விபரங்கள் என்றும் மேலும் இரண்டு வகைகளாகப் புள்ளி விபரங்கள் பிரிக்கப்படுகின்றன. உதாரணத்திற்கு மதிப்பெண், எடை, உயரம், மனிதர்கள் எண்ணிக்கை தொடர்பான விபரங்களை எண்களால் கூற இயலும். ஆனால் ஒருவருடைய மதம், சாதி, மொழி, நேர்மை, திறமை போன்ற இன்னும் சில குணங்களுக்கு எண்கள் கொடுக்க இயலாது. (அப்படி அவற்றிற்கு எண்கள் கொடுக்கப்பட்டால் அவை குழப்பங்களை கூட்டுமேயொழிய குழப்பங்களைக் களையாது).

சில செய்திகளை முழு எண்களால் மட்டுமே அளவிடவும் கூறவும் முடியும். பின்ன எண்களால் குறிப்பிட முடியாது. உதாரணத்திற்கு குழந்தைகள், பேனாக்கள், தொலைக்காட்சிப் பெட்டிகள், வீடுகள் (உ.ம். 10, 15, 20) போன்றவற்றை முழு எண்களால்தான் குறிப்பிட முடியும். ஆனால், சில புள்ளி விபரங்களை இன்னும் மிகத் துல்லியமாக

பின்ன எண்களாலும் குறிப்பிடலாம். உதாரணம் வயது, உயரம், எடை, மதிப்பெண்கள் (உ.ம். 1.5, 2.3, 4.1, 5.2) போன்றவை. எனவே முழு எண் விபரங்கள் என்றும் பின்னமாக வருபவற்றைத் தொடர் எண் விபரங்கள் என்றும் அழைக்கலாம்.

ஓர் ஆய்வுக்கான புள்ளி விபரங்களைச் சேகரிக்கும் முன்பு எப்படிப்பட்ட புள்ளி விபரங்கள் வேண்டும் என்று முடிவு செய்வது நல்லது. ஒரே ஆய்வுக்குப் பலதரப்பட்ட புள்ளி விபரங்களும் தேவைப்படலாம்.

மொத்தம் மற்றும் மாதிரி

ஓர் ஆய்வுக்குட்படுத்தப்படும் புள்ளி விபரங்களை எந்த அளவுக்கு விரிவாகச் சேகரிக்க வேண்டும் என்பதும் ஒரு முக்கியமான முடிவாகும். ஆய்வுக்குட்படுத்தப்படும் அனைத்து உறுப்புக்களிலிருந்தும் (elements) புள்ளி விபரங்கள் சேகரிக்கப்பட்டால் அது முழு அல்லது மொத்த விவரணம் என்றும் சில கூறுகளிடமிருந்து புள்ளி விபரங்கள் சேகரிக்கப்பட்டால் அது மாதிரி விவரணம் (கூறு, Sample survey) என்றும் அழைக்கப்படுகிறது. அனைத்து உறுப்புக்கள் பற்றியும் புள்ளி விபரங்கள் சேகரித்தால் ஆய்வு மிகச் சரியாக இருக்க வாய்ப்பிருக்கிறது என்று நம்பப்படுகிறது. ஆய்வின் முடிவுகள் மிகத் துல்லியமாக இருக்கவும் அந்த முடிவுகளைக் கொண்டு கணிக்கப்படும் எதிர்கால விபரங்களும் சரியாக இருக்கவும் வாய்ப்புள்ளது. எனவே முழுமையான கணக்கிடுதல் மிக நேர்த்தியான ஒன்றாகக் கருதப்படுகிறது. எனினும் முழுமைக் கணிப்பிலும் சில சிரமங்கள் உள்ளன. பல சமயம் முழுமையான மொத்தக் கூறுகளையும் ஆய்வுக்குட்படுத்த முடியாத சூழ்நிலையிருக்கலாம். உதாரணமாக ஒரு நிறுவனத்தால் உற்பத்தி செய்யப்படும் அனைத்துக் குழல் விளக்குகளையும் சோதனை செய்து பார்த்து விட்டால் விற்பதற்கு விளக்கு இருக்காது. அதுபோல ஒரு

பானையில் உள்ள எல்லா சோற்றுப் பருக்கைகளையும் வெந்து விட்டதா என்று பார்த்தால் சாப்பிடச் சோறு இருக்காது. எனவே அப்படிப்பட்ட சூழ்நிலைகளில் சில உறுப்புகளை ஆய்தலே நலம்.

சில குழல் விளக்குகளை எடுத்து ஆய்வு செய்து அந்த முடிவை வைத்து அனைத்துக் குழல் விளக்குகளும் எத்தனை மணி நேரம் ஒளி தரும் என்று கூறுவதே மிகப் பொருத்தமாகும். பல சமயங்களில் முழுமையாக எல்லா உறுப்புக்களையும் ஆய்வுக்குட்படுத்துவது என்பது சிரமமான செயலாகவும் இருக்கலாம்; கால தாமதமாகலாம்; பணச் செலவு அதிகமாக ஆகலாம்; அதிக எண்ணிக்கை கொண்ட உறுப்புக்களை ஆழ்ந்த ஆய்வுக்கு உட்படுத்தும்போது பல இடங்களில் தவறுகள் நேரலாம். இதனாலேயே சிறிய அளவிலான பொறுத்தமான உறுப்புக்களை எடுத்து ஆழ்ந்த ஆய்வுக்கு உட்படுத்துவதே விரும்பப்படுகிறது. இப்படிச் செய்வதால் மிகக் கவனமாக ஆய்வுகள் செய்ய வாய்ப்பிருக்கிறது; கவனக்குறைவால் நிகழும் தவறுகளைக் கணிசமாகக் குறைத்து விடலாம். எனவேதான் பெரிய அளவிலான விரிவான ஆய்வுகளில் மொத்தத்தின் சில பகுதிகள் மட்டும் மாதிரியாகத் தேர்ந்தெடுக்கப்பட்டு ஆய்வு செய்யப்படுகின்றன. இந்தியாவைப் பொறுத்தமட்டில் மக்கள்தொகை பத்தாண்டுகளுக்கு ஒருமுறை முழுமையான ஆய்வுக்கு உட்படுத்தப்படுகிறது. அனேகமாக மற்ற எல்லா விபரங்களும் பகுதிகளாகவே எடுத்து ஆயப்படுகின்றன. இதற்கு உதாரணமாக தேசிய மாதிரி ஆய்வு நிறுவனத்தைக் (National Sample Survey Organisation) கூறலாம்.

மாதிரிகளின் (Samples) வகைகள்

ஒரு மொத்தத்திலிருந்து சில பகுதிகளை (elements) மட்டும் எடுத்து ஆய்வு செய்வது இன்று பரவலாகக் காணப்படுகிறது. ஆனாலும் ஒரு மொத்தத்தின் சில

உறுப்புக்களைத் (elements) தேர்ந்தெடுக்கும்போது சரியான முறையில் அதனைச் செய்தால்தான் துல்லியமான பயன்படுத்தக் கூடிய முடிவுகள் கிடைக்கும். அப்படியில்லை எனில் கிடைக்கின்ற முடிவுகள் எந்தவிதப் பயனும் இல்லாமல் போய்விடும்; அம்முடிவுகளைப் பயன்படுத்தினாலும் அவை தவறான விளைவுகளையே தரும். ஒரு முழுமையிலிருந்து (Population, universe) சில பகுதிகளை மட்டும் தேர்ந்தெடுக்க பல முறைகள் உள்ளன. பொருத்தமான முறையைக் கீழ்க்காணும் கருத்துக்களை மனத்தில் கொண்டு தெரிந்தெடுக்க வேண்டும்.

1. ஆய்விற்கு எடுக்கப்பட்டுள்ள பொருளின் தன்மை
2. உறுப்புக்களைத் தேர்ந்தெடுக்கப்பெற வேண்டிய முழுமைத் தொகுதியின் தன்மை
3. மாதிரிக்கூறை ஆய்ந்து பெறும் முடிவுகள் எவற்றிற்குப் பயன்படுத்தப்பட போகின்றனவோ அவற்றின் தன்மைகள்

மாதிரிகளைச் சிறிய மாதிரிகள் (Small samples) பெரிய மாதிரிகள் என்று இரண்டு வகைகளாகப் பிரிக்கலாம். ஒரு முழுமைத் தொகுதியிலிருந்து எடுக்கப்பட்ட கூறுகளின் எண்ணிக்கை 30க்கும் குறைவாக இருந்தால் அந்த மாதிரி சிறிய மாதிரி எனவும், 30 அல்லது அதற்குப் பெரியதாக இருந்தால் பெரிய மாதிரி எனவும் அழைக்கப்படுகின்றன. சூழ்நிலையைப் பொறுத்து உறுப்புக்களின் எண்ணிக்கையை முடிவு செய்ய வேண்டும். நிறைய உறுப்புக்களை ஆய்வு செய்ய முடியாத நேரங்களில் சிறிய மாதிரியையும் மற்ற நேரங்களில் பெரிய மாதிரியையும் தேர்ந்தெடுத்துக் கொள்ளலாம். அதனைப் பொறுத்தே திட்டப்பிழை காண்பதற்கான (இது பற்றி பின்னர் விரிவாக காணலாம்) உபாயங்களும் அமையும்; எடுகோள்கள் சோதனை செய்யப்படும்போது (இது பற்றியும் பின்னர் விரிவாகக் காணலாம்) சோதனை முறைகளிலும் வித்தியாசங்கள் காணப்படும்.

மேலும் ஒரு முழுமைத் தொகுதியிலிருந்து உறுப்புக்களைத் தெரிவு செய்யும் முறையை வைத்தும் மாதிரிகளை இருவகைகளாகப் பிரிக்கலாம். ஒரு முழுமைத் தொகுதியில் உள்ள ஒவ்வொரு உறுப்புக்கும் சம வாய்ப்பளிக்கப்பட்டு சில உறுப்புக்கள் மட்டும் தேர்ந்தெடுக்கப்பட்டால் அம்முறை இயைபிலா அல்லது நிகழ்தகவு மாதிரி முறை (random or probability sampling) என அழைக்கப்படுகிறது. அப்படியின்றி, சில குறிப்பிட்ட காரணங்களுக்காக அல்லது நோக்கங்களுடன் சில உறுப்புக்கள் தேர்ந்தெடுக்கப்பட்டால் அம்முறை காரண மாதிரி முறை (Purposive, deliberate or judgement sampling) என அழைக்கப்படுகிறது. ஒரே ஆய்வுக்கே தேவைக்குத் தகுந்தாற்போல் மேற்கூறிய இரண்டு முறைகளையும் பயன்படுத்தியும் உறுப்புக்கள் தேர்ந்தெடுக்கப்படலாம். உதாரணத்திற்கு, ஒரு மாநிலத்தில் உள்ள சில மாவட்டங்களைக் காரண மாதிரி அடிப்படையில் தெரிவு செய்து அங்குள்ள கிராமங்களை இயைபிலா மாதிரி முறையில் தெரிவு செய்யலாம். அல்லது ஒரு மாவட்டத்தில் உள்ள கிராமங்களைக் காரண அடிப்படையில் தேர்ந்தெடுத்துவிட்டு அந்தக் கிராமங்களில் உள்ள வீடுகளை இயைபிலா மாதிரி முறையில் தெரிவு செய்யலாம்.

ஒரு மாதிரியில் (sample) பல உறுப்புக்கள் (elements, or number of observation) தெரிந்தெடுக்கப்படலாம். அதேபோல, ஒரு முழுமைத் தொகுதியிலிருந்து பல மாதிரிகள் தெரிவு செய்யப்படலாம். அந்த மாதிரிகளில் உள்ள உறுப்புக்களின் எண்ணிக்கையிலும் மாறுபாடு இருக்கலாம். ஒரு முழுமைத் தொகுதிக்கு ஒரு மாதிரி என்னும் அடிப்படையில் பல முழுமைத் தொகுதிகளிலிருந்து பல மாதிரிகளும் தெரிவு செய்யப்படலாம். முழுமைத் தொகுதியின் பரவல் தன்மையைப் பொறுத்தும், அங்குள்ள உறுப்புகளின்

தன்மைகளைப் பொறுத்தும், முழுமைத் தொகுதியில் உள்ள உறுப்புக்களின் எண்ணிக்கையும் அமையும். உறுப்புக்களை முழுமைத் தொகுதியிலிருந்து எடுக்கும் விதத்தைப் பொறுத்தும் ஆய்வுக்குட்படுத்தப்பட்ட உறுப்புக்களே மீண்டும் மீண்டும் வருவதற்கு வாய்ப்புக் கொடுத்தும் (with replacement) வாய்ப்பு கொடுக்காமலும் (without replacement) உறுப்புக்களைத் தெரிவு செய்யலாம். இந்த தெரிவு செய்யும் முறைகளைப் பொறுத்து திட்டப்பிழைகள் கணிக்கும் முறைகளிலும் வேறுபாடுகள் இருக்கும்.

இயைபிலா மாதிரி முறையை (random sampling) மீண்டும் இருவகைகளாகப் பிரிக்கலாம்.

1. நிபந்தனையற்ற எளிய இயைபிலா அல்லது சமவாய்ப்பு மாதிரிகள் (unrestricted simple random samples)
2. நிபந்தனையுடைய இயைபிலா அல்லது சமவாய்ப்பு மாதிரிகள் (restricted random samples)

நிபந்தனையற்ற எளிய மாதிரிகள் கீழ்க்காணும் இரு முறைகளில் ஏதேனும் ஒரு முறையைப் பின்பற்றித் தேர்ந்தெடுக்கப்படுகின்றன.

1. சீட்டுக்குலுக்கல் முறை (Lottery Method)
2. இயைபிலா மாதிரி அடிப்படையில் அமைக்கப்பட்டுள்ள எண்களைப் பயன்படுத்தும் முறை (random numbers method)

நிபந்தனையுடைய இயைபிலா மாதிரிகளை மூன்று வகைகளாகப் பிரிக்கலாம்.

1. பிரிவுபடுத்தப் பெற்ற மாதிரி (Stratified sampling)
2. திரண்ட மாதிரி (Cluster sampling)
3. முறையாகத் தேர்ந்தெடுக்கப்படும் மாதிரி (systematic sampling)

ஆய்வுச் சூழல்களுக்கேற்ப மேற்காணும் முறைகளில் ஒன்றையோ ஒன்றுக்கு மேற்பட்ட முறைகளையோ பயன்படுத்திக் கொள்ளலாம். அவ்வாறு மாதிரி முறைகளைத் தேர்ந்தெடுக்கும் பொழுது, முழுமை (universe)யின் பரவல் தன்மை (Pattern of distribution), முழுமையின் அளவு (size of population), எடுக்கப் போகின்ற மாதிரியின் அளவு (sample size), கூறுகளுக்கிடையே உள்ள உறவு (relationship betwen the elements) போன்றவைகளை மிகக்கவனமாக அறிந்து கொள்ளுதல் அவசியம். இவ்வாறாகத் தேர்வு செய்யப்பட்ட மாதிரிகளின் குணங்கள் (Statistics) முழுமையின் குணங்களோடு (Parameters) பொருந்திப் போவனவாக அமைய வேண்டும். அப்படியில்லையெனில் மாதிரி மூலம் கிடைக்கப்பெறுகின்ற ஆய்வு முடிவுகள் பயனற்றுப் போய்விடும்; அப்படிப்பட்ட ஆய்வு முடிவுகளை வைத்து திட்டங்களையோ கொள்கைகளையோ வகுப்பதோ, திட்டங்களின் விளைவுகளை ஆய்வதோ சரியில்லாமல் போய்விடலாம். சரியான மாதிரிகளைத் தேர்ந்தெடுத்த பின்னர் தான் புள்ளி விபரங்களைச் சேகரிக்க முயற்சிக்க வேண்டும்.

எப்படி மாதிரி எடுக்க வேண்டும் என்பதும் முக்கியம். உதாரணத்திற்கு, இரண்டு பெண்கள், மூன்று பெண்கள் உள்ள குடும்பங்கள் எத்தனை என்று அறிய பெண்கள் பயிலும் ஒரு கல்லூரியை எடுத்து அங்குள்ள மாணவிகளிடம் கேட்டு அறிவது தவறான வழியாகலாம். ஏனெனில் சில மாணவிகளின் சகோதரிகளும் அதே கல்லூரியில், சில வசதிக்காக, படிக்கலாம். அப்படியிருக்க வாய்ப்பு உள்ளது. அப்படியானால் ஒரே குடும்பத்திலிருந்து வரும் மூன்று சகோதரிகளும் எங்கள் வீட்டில் மூன்று பெண்கள் படிக்கிறோம் என்று சொல்லலாம். இது வெவ்வேறு வகுப்பறைகளில் நடந்திருக்கலாம். அப்படியானால், ஆய்வு செய்பவர் மூன்று பெண்கள் உள்ளது மூன்று குடும்பங்கள் என எடுத்துக் கொள்வார். ஆனால், அங்கு மூன்று பெண்கள் உள்ளது ஒரே ஒரு குடும்பம்தான்.

இவ்வாறாக எண்ணிக்கை தவறாகி விடலாம். எனவே, இந்த மாதிரி சரியில்லாமல் போகலாம். இப்படிப்பட்ட ஆய்வுக்கு கல்லூரியையோ, பள்ளியையோ தேர்ந்தெடுக்காமல், வீடுகளை எடுத்து அங்கிருந்து இப்படிப்பட்ட புள்ளி விபரங்களைச் சேகரித்தால் மேலே கூறிய பிரச்சனையைத் தவிர்த்து விடலாம்.

புள்ளி விபர நிறுவனங்கள்

புள்ளி விபரங்களைப் பெறுவதற்கு இன்று பல வழிகள் உள்ளன. அரசு மற்றும் தனியார் நிறுவனங்களும், ஆய்வகங்களும் பலவிதமான இதழ்கள், குறிப்புக்கள் மற்றும் அறிக்கைகளை வெளியிட்டு வருகின்றன. அவற்றின் மூலம் பொருத்தமான புள்ளி விபரங்களைப் பெற முயற்சிக்கலாம். இவை ஓர் ஆய்வுக்கான அடிப்படை உண்மைகளைப் பற்றித் தெரிந்து கொள்ள உதவலாம். ஓர் ஆய்வுக்கான சரியான அமைப்பை உருவாக்குவதற்கும், சரியான ஆய்வு அலகினைத் தேர்ந்து கொள்ளவும் மேலே குறிப்பிட்ட விபரங்கள் உதவலாம். ஆனாலும் பல சமயங்களில் அந்தப் புள்ளி விபரங்களைத் தேவைக்கேற்ப சீரமைக்க வேண்டியும் வரலாம். இவ்வாறான புள்ளி விபரங்களைச் சேகரிப்பதற்கும், ஆய்வுக்கு அளிப்பதற்கும் உள்ள அரசு நிறுவனங்களில் மைய வங்கி (Reserve Bank of India) தேசிய மாதிரி ஆய்வு நிறுவனம் (National Sample Survey Organisation) போன்றவைகளைக் கூறலாம். ஒவ்வொரு வருவாய்க் கிராமத்தின் (Revenue village) மக்கள் தொகை பற்றிய விபரங்களும் இந்தியாவில் நூறு ஆண்டுகளுக்கும் மேலாகக் கிடைக்கின்றன. மாவட்ட மக்கள் தொகைக் கையேடுகளில் (District Census Hand Book) நிலங்களின் வகைகளைப் பற்றிய புள்ளி விபரங்களும் மக்களைப் பற்றிய புள்ளி விபரங்களும் கிடைக்கின்றன. கால்நடைகள் மற்றும் பொருளாதார நிறுவனங்கள் பற்றிய புள்ளி விபரங்களும் தற்போது

நிறையக் கிடைக்கின்றன. தனியார் துறையில் மும்பேயில் தலைமையிடம் கொண்டுள்ள இந்தியப் பொருளாதாரத்தைக் கண்காணிக்கும் மையம் (Centre for Monitoring Indian Economy) மிக நேர்த்தியான மிக அதிகமானப் புள்ளி விபரங்களைத் தருகின்றது. சமீப காலமாக தொழில் நிறுவனங்கள் பற்றிய ஏராளமான புள்ளி விபரங்களைக் கணினி உதவி கொண்டு பெற்றுக் கொள்ளும் வகையில் ப்ரோவஸ் (Prowess) என்ற சிப்பமும் (package) கிடைக்கின்றது.

முதல்நிலைப் புள்ளி விபரங்கள்

எவ்வளவோ இரண்டாம் நிலைப் புள்ளி விபரங்கள் தற்போது கிடைத்த போதும் பல நிலைகளில் தேவையான புள்ளி விபரங்கள் தேவையான வகையில் கிடைக்காமலும் போகின்றன. அப்படிப்பட்ட சூழ்நிலைகளில் ஆய்வாளர் தாமே சரியான புள்ளி விபரங்களைச் சேகரிக்க வேண்டிய நிலைமையும் வந்துவிடுகிறது. அப்படி வரும்போது, மறுபடியும் வசதிக்கேற்பவும் அவசியத்திற்கேற்பவும் பலவாறாக முதல்நிலைப் புள்ளிவிபரங்கள் சேகரிக்கப் படுகின்றன. கீழ்வரும் நான்கு முறைகள் மிகப் பலரால் பயன்படுத்தப்படுகின்றன.

1. நேரடியாக ஆய்வாளரே சேகரிக்கும் முறை (Direct Personal Investigation / Interview)
2. உள்ளூர் நிருபர் மூலம் பெறும் முறை
3. தபால் வழி வினாத் தொகுதி முறை
4. கணிப்பாளர் மூலம் பட்டியல் முறை

உண்மையான புள்ளி விபரங்களைச் சேகரிப்பது ஒரு கடினமான பணி. நேரடியாக ஆய்வாளரே புள்ளிவிபரங்கள் சேகரிக்கும் முறையே சிறந்த முறை. ஆனாலும் இதையே பலர் பலவிதமாகச் செயல்படுத்தி வருகின்றனர். யாரைப் பற்றிப் படிக்கின்றோமோ அவரிடமே புள்ளி விபங்களைச் சேகரிக்கலாம். அப்படிச் செய்யும்போது சிரமங்கள்

இருந்தாலும் எக்காரணத்திற்கு அந்த ஆய்வு செய்யப்படுகின்றதோ அதை விளக்கமாகத் தெளிவுபடுத்தி விட வேண்டும். அவ்வாறின்றி வேறு யாதேனும் காரணங்களைக் கூறினால் புள்ளி விபரங்கள் தருபவரும் அதற்குத் தகுந்தாற்போல் புள்ளி விபரங்களை மாற்றி உண்மையை மறைத்து விடுவார். நாம் யாரைப் பற்றிப் படிக்கின்றோமோ அவர் செய்திகளைத் தர விரும்பாமலோ அல்லது தரக்கூடிய நிலையில் இல்லாமலோ (மது அருந்தி விட்டு) இருக்கலாம். அப்படிப்பட்ட சூழ்நிலையில் மற்றவர்களிடமிருந்து செய்திகளைத் திரட்டலாம். ஆனாலும் அச்செய்திகளைப் பற்றி வேறு சிலரிடமும் கேட்டு உறுதிப்படுத்திக் கொள்வதே நல்லது. சில சமயம் பொதுவானவை பற்றிச் செய்திகள் சேகரிக்க அந்த செய்திகளைப் பொறுத்தமட்டில் யார் யார் நிறையச் செய்தி வைத்திருப்பார்கள் என்று நம்பப்படுகிறார்களோ அவர்களை அணுகி செய்திகளைச் சேகரிக்கலாம். அவர்களை ஒரு குழுவாக அமைத்துக் கலந்துரையாடி (Focussed Group Discussion - FGD)யும் செய்திகளைச் சேகரிக்கலாம். ஆய்வுக்கான உண்மைக் காரணத்தைக் கூறினாலும் உண்மையான செய்திகள் கிடைக்காமல் போகலாம் என்று நினைப்பவர்கள் (மிகச் சிலரே) எப்படிப்பட்ட மக்களைப் பற்றி ஆய்வு செய்ய நினைக்கிறார்களோ அவர்களுடன் தாமும் மற்றவர்களுக்குச் சந்தேகம் வராமல் ஓர் அங்கத்தினராக மாறி (உண்மையைச் சொல்லாமல்) செய்திகள் திரட்டுகின்றனர். இம்முறை மூலம் உண்மை வெளிவர வாய்ப்புண்டு. இது கடினமான செயல். ஆனாலும் சிலர் பிச்சைக்காரராக மாறியும், சிலர் வாகனம் பழுதுபார்க்கும் நிறுவனத்தில் ஒரு தொழிலாளியாக மாறியும், சிலர் உணவு விடுதியில் பணியாளாக மாறியும் அவர்களைப் பற்றிய உண்மையான அதிக செய்திகளைத் திரட்டியுள்ளனர். இப்படிப் பங்கு கொண்டு செய்திகள் திரட்டுவது (Participant

observation), பங்கு கொள்ளாமல் (Non-participant) செய்திகள் திரட்டுவதை விட நல்லது. ஆனாலும், சிரமம் கருதி பங்கு கொண்டு செய்தி திரட்டும் முறையைப் பின்பற்றுபவர்கள் மிகச் சிலரே.

மேலே கூறப்பட்டுள்ள முறைகள் ஒவ்வொன்றும் தனிச்சிறப்பையும் தனிச் சிரமங்களையும் கொண்டுள்ளது. ஒரே ஆய்வுக்கே பலதரப்பட்ட முறைகளும் பயன்படுத்தப்படுகின்றன. ஆய்வாளர் நேரடியாகச் சென்று செய்திகள் சேகரிக்காவிட்டால் அப்படிப்பட்ட செய்திகள் மிகப் பெரிய சந்தேகத்தை உள்ளடக்கியதாகவே அமையும். செய்திகள் சேகரிக்க இருப்பவர்களுக்குச் சரியான பயிற்சி அளிப்பது அவசியம். அப்படியே பயிற்சி அளித்தாலும், சில கருத்துருக்களை சரியாகப் புரியாமலேயே புள்ளி விபரங்கள் சேகரிக்க முற்படுபவர்களும் உண்டு. அப்படிப்பட்ட புள்ளி விபரங்கள் அவ்வளவாக ஆய்வுக்குப் பயன்படாமலேயே போகவும் வாய்ப்பு உள்ளது. எனவே, எல்லா நிலைகளிலும் ஆய்வாளர் மிகக் கவனமாக இருப்பது அவசியம்.

புள்ளி விபரங்கள் சேகரித்தல்

எதனைப் பற்றி யாரிடம் எவ்விதப் புள்ளி விபரங்கள் எவ்விதம் சேகரிக்கப்பட இருக்கின்றன என்ற முடிவுகளைத் தெளிவாக எடுத்துக்கொண்ட பிறகு புள்ளி விபரங்கள் சேகரிக்க ஆரம்பிக்கலாம். அப்பொழுது அனைத்து நோக்கங்களும் சரியாகத் தெரிந்திருக்கப்பட வேண்டும். கருத்துருக்கள் சரியாகப் புரிந்திருக்கப்பட வேண்டும். கேள்விகள் சரியாக வரிசைப்படுத்தப்படவேண்டும். முன்னுக்குப் பின் முரணான கேள்விகள் தவிர்க்கப்பட வேண்டும். பதில் சொல்ல இருப்பவரின் மனத்தைப் புண்படுத்தக் கூடியவையாகவோ, எரிச்சலூட்டக்கூடியவையாகவோ கேள்விகள் அமையக் கூடாது. புள்ளி விபரங்கள் சேகரிக்கும்போது அவசரம் காட்டுதலோ பதற்றப்படுதலோ கூடாது. பதில் சொல்ல

இருப்பவரின் மனநிலையைப் புரிந்து அவரிடம் சரியான ஒப்புதல் பெற்று உண்மையான காரணத்தைக் கூறி பொறுமையோடு வினாக்களை முன்வைக்க வேண்டும். இந்நிலையில் பலரும் பலவிதமான தவறுகள் செய்ய வாய்ப்புள்ளது. அதிகமான விவரங்களை ஒரே நேரத்தில் சேகரிக்க முற்படுவதோ அல்லது போதுமான விபரங்களைச் சேகரிக்காமல் விட்டு விடுவதோ நல்லதல்ல. இத்தவறினைச் செய்யாமலிருக்க பல வழிகள் உள்ளன. எடுத்துக்கொண்ட ஆய்வுத் தலைப்புக்குத் தொடர்பான பல புத்தகங்களையும் கட்டுரைகளையும் படிக்க வேண்டும். மேலும், அதே மாதிரி அனுபவம் கொண்ட ஆய்வாளர்களிடம் கலந்தாலோசனை செய்யவும் வேண்டும். இது தொடர்பான இரண்டாம் நிலைப் புள்ளி விபரங்களை நன்றாகப் படித்து ஆராய்ந்து விட்டு முதல் நிலைப்புள்ளி விபரங்களைச் சேகரிப்பது நல்லது. இப்படித்தான் நல்ல ஆய்வாளர்கள் செய்வார்கள்.

புள்ளி விபரங்களைத் தவறாகப் பயன்படுத்துதலைத் தவிர்த்தல்

புள்ளி விபரங்களைப் பயன்படுத்தித் தவறான அபிப்ராயங்களையும் உருவாக்கலாம். பல சமயங்களில் கிடைத்துள்ள ஒரேவகையான புள்ளிவிபரங்களிலிருந்து அவற்றைச் சாதுரியமாகப் பயன்படுத்தி பல வேறுபாடான கருத்துக்களையும் முடிவுகளையும் கூடக் காட்டமுடியும். உதாரணத்திற்கு, ஒரு கணினி விற்பனையாளர் தன்னுடைய நிறுவனம் உற்பத்தி செய்யும் கணினியை 80 விழுக்காடு மக்கள் பயன்படுத்துகின்றனர் என்பார். ஆனால் உண்மை என்ன என்றால், கணினி பயன்படுத்தும் மக்களில் (உதாரணத்திற்கு 10 விழுக்காடு மக்கள்) 80 விழுக்காடு கணினி உபயோகிப்பாளர்கள் அவருடைய கணினியினைப் பயன்படுத்துகின்றனர் என்பதுதான். அதாவது 8 விழுக்காடு மக்கள்தான் அவருடைய கணினியினைப் பயன்படுத்துகின்றனர் என்பதே அதன் பொருள்.

தன்னுடைய அரசு இந்த ஆண்டு 12 கோடி ரூபாய் ஓரினத்தைச் சார்ந்த மக்களுக்காகச் செலவு செய்கிறது என்று ஓர் இந்திய அரசியல்வாதி கூறி தான் அவ்வின மக்களுக்கு அதிக நன்மை செய்வதாகக் கூறுவார். மற்றொருவர் அந்தப் பணத்தை நபர் ஒருவருக்குக் கணக்கிட்டு (அந்தத் தொகையை மக்கள் தொகையால் வகுத்து) பாருங்கள் இந்த அரசு ஒரு நபருக்கு ஓர் ஆண்டுக்கு ஒரு ரூபாய் தான் செலவு செய்கிறது; இது ஒரு பெரிய தொகையா என்று கேட்டு அவர் மக்களின் ஆதரவைச் சேர்ப்பார். மொத்தமாகச் சொல்லும்போது ஒரு பொருள் தெரியும்; அதையே ஒரு நபருக்குக் கணக்கிட்டுச் சொன்னால், முன்கூறியதற்கு நேர் எதிரான விளைவு தரும். சராசரிகளையும், வளர்ச்சி வீதங்களையும் வைத்தும் பலவேறு முரண்பாடான முடிவுகளைக் கொண்டு வரலாம். உதாரணத்திற்கு ஒரு மாணவன் நான் போன தேர்வில் பெற்ற மதிப்பெண்ணைவிட இந்தத் தேர்வில் 100 சதவீதம் அதிகம் வாங்கியுள்ளேன்; ஆனால் என் தம்பி 40 சதவீதம்தான் அதிகம் பெற்றுள்ளான். எனவே நான்தான் திறமையானவன் என்பான். அவனுடைய தம்பியோ, என் அண்ணன் முதலில் வாங்கியதைவிட இப்பொழுது 10 மதிப்பெண்கள்தான் அதிகம் பெற்றுள்ளான்; நானோ 20 மதிப்பெண்கள் அதிகம் பெற்றுள்ளேன். எனவே நான்தான் திறமைசாலி என்பான். அவர்கள் முதல் தேர்வில் பெற்ற மதிப்பெண்கள் தெரியா விட்டால் குழப்பம் வரலாம். அண்ணன் முதலில் பெற்ற மதிப்பெண் 100க்கு 10; இரண்டாவதாகப் பெற்ற மதிப்பெண் 100க்கு 20; எனவே 100 விழுக்காடு அதிக மதிப்பெண்கள். தம்பி முதலில் பெற்ற மதிப்பெண் 100க்கு 50, இரண்டாவதாகப் பெற்ற மதிப்பெண் 100க்கு 70; எனவே (50க்கு 20 கூட என்பது) 40 விழுக்காடுதான் அதிகம்; ஆனாலும் இதுதான் கடினம்.

இதுபோல சில சமயம் மொத்த மதிப்பெண்களைச் சரியாகக் கூறாமலும் சிலர் குழப்பம் விளைவிக்கலாம்.

உதாரணத்திற்கு நான் 80 மதிப்பெண் பெற்றுள்ளேன். என் தம்பி 70 மதிப்பெண்கள்தான் பெற்றுள்ளான் என்று சொல்லி அண்ணன் நல்ல பேர் வாங்கிக் கொள்ள முயற்சிப்பான். தம்பி, நான் 100க்கு 70 மதிப்பெண்கள் பெற்றுள்ளேன்; என் அண்ணனோ 200க்கு 80 மதிப்பெண்கள் தான் பெற்றுள்ளான் என்பான். இதுபோன்ற சூழ்நிலைகளில் சதவீதத்தில் (தம்பி 70 விழுக்காடு அண்ணன் 40 விழுக்காடு என்று) கூறுவது குழப்பத்தைக் குறைக்கும்.

ஒவ்வொரு தாளிலும் 100 மதிப்பெண்களாகவும் தேர்ச்சி பெற 40 மதிப்பெண்கள் தேவை என்ற நிலையும் உள்ளபோது மூன்று மாணவர்கள் பெற்றுள்ள மதிப்பெண்கள் கீழேதரப்படுகின்றன.

	முதல் தாள்	இரண்டாம் தாள்	மூன்றாம் தாள்	நான்காம் தாள்	மொத்தம்
மாணவர் 1	20	50	60	70	200
மாணவர் 2	70	60	50	30	210
மாணவர் 3	45	45	45	45	180

இப்பொழுது முதல் இரு மாணவர்களும் தலா ஒரு தாளில் தோல்வியடைந்துள்ளனர். ஆனால் இருவருமே மூன்றாம்மாணவரைவிடஅதிக மதிப்பெண்கள் பெற்றுள்ளனர்.

இருப்பினும் மூன்றாம் மாணவர் நான்கு தாள்களிலும் தேர்ச்சி பெற்றுள்ளார். எனவே மூன்றாமவர் சிறந்த செயல்திறனுடையவர் எனலாம். முதல் இரு மாணவர்களுள் இரண்டாம் மாணவர் முதல் மாணவரின் மதிப்பெண்களைவிட (200) அதிக மதிப்பெண்கள் (210) பெற்றுள்ளார். எனவே திறமைசாலி எனலாம். ஆனாலும் அவரின் செயல்திறன் ஒவ்வொரு தாளிலும் குறைந்து கொண்டே சென்றுள்ளது. முதலாமவரின் மதிப்பெண்கள் கூடிக்கொண்டே சென்றுள்ளன. இப்படிப்பட்ட சூழ்நிலைகளில் புள்ளியியல் தரக்கூடிய

மதிப்புக்கள் அல்லது அளவீடுகள் (சராசரி, திட்டவிலக்கம் போன்றவை) முழு விளக்கத்தையும் தர இயலாது. எனவே, இயல்பான புள்ளி விபரங்களை வைத்து சூழ்நிலைக்குத் தேவையான மற்றும் பொருத்தமான முடிவுகளை எடுப்பதே சிறந்ததாகும். மேலே கூறப்பட்டுள்ள உதாரணத்திற்கு, வளர்ச்சி வீதம் கண்டுபிடித்திருந்தால் அது முதல் மாணவருக்குக் கூட்டல் குறியோடும் இரண்டாம் மாணவருக்குக் கழித்தல் குறியோடும் வந்திருக்கும்; அதன் மூலம் இரண்டாம் மாணவர் மோசமாகியிருக்கிறார் என்றும் முதல் மாணவர் முன்னேறி இருக்கிறார் என்றும் கூறியிருக்கலாம்.

புள்ளி விபரங்களை ஒழுங்குபடுத்துதல்
(Organisation of Statitstics)

புள்ளியியல் முறைகளில் இரண்டாவது நிலையாக புள்ளிவிபரங்களை ஒழுங்குபடுத்துதல் உள்ளது. (புள்ளி விவரங்களைச் சேகரித்தல் முதல்நிலை ஆகும்). சேகரிக்கப்பட்ட புள்ளி விவரங்கள் ஆய்வுக்குப் பொருத்தமாக ஆக்கப்படுகின்றன. புள்ளி விவரங்களைப் பாகுபடுத்தியும் (Classification) பட்டியலிட்டும் (Tabulation) அவற்றை ஆய்வு செய்வதற்குப் பொருத்தமாக அமைக்கலாம்.

சேகரிக்கப்பட்ட புள்ளி விபரங்களை அவற்றின் தன்மைக்கேற்ப முதலில் பாகுபடுத்திப் பிரிக்க வேண்டும். எடுத்துக்காட்டாக ஒரு கிராமத்தில் உள்ள வீடுகளிலிருந்து சேகரிக்கப்பட்ட விவரங்களை முதன்மைப்பட்டியலில் (Master Table) பதிவு செய்யும் முன்னர் அவ்விவரங்களைப் பிரித்துக் கொள்வது நல்லது. அப்படிப் பிரிப்பதற்கு ஆய்வின் நோக்கங்களும் (Objectives) எடுகோள்களும் (Hypotheses) வழிகாட்டுதல்களாக அமையும். உதாரணத்திற்கு, ஒரு கிராம அமைப்பையும் செயல்பாட்டு விதத்தையும் நிர்ணயிப்பதில் மிக முக்கிய பங்கு வகிப்பது சமூகப் பிரிவுகளா (Caste System)

அல்லது நிலச் சொந்தம் அடிப்படையிலான வகுப்புப் பிரிவுகளா (land ownership based classes) என்று காண்பது ஆய்வின் நோக்கமானால், அக்கிராமத்திலிருந்துசேகரிக்கப்பட்ட விவரங்களை மேற்கூறியவற்றின் அடிப்படையில் பிரித்து முதன்மைப் பட்டியல் செய்தால் பிறகு அதிக நேரம் விரயம் ஆகாமல் தடுக்கலாம். எனவே, மதம், சாதி வாரியாக வீடுகளைப் பிரித்துவிட்டுப் பிறகு ஒவ்வொரு மதம், சாதியில் உள்ள வீடுகளையும் சொந்தமாக உள்ள நிலத்தின் அடிப்படையில் பிரித்தால், பிறகு மதம் சாதி வாரியாகவும் வகுப்பு வாரியாகவும் விவரங்களைத் திரட்டுவது எளிதாகி விடும். இல்லையேல், சாதி, மதம், வகுப்பு வாரியாக முதன்மைப்பட்டியலிலிருந்து தேடிக் கண்டுபிடிப்பதில் அதிக நேரம் வீணாகலாம்.

மாணவர்களைப் பற்றிய ஆய்வு என்றால் முதலிலேயே மாணவர்களை ஆண்கள் பெண்கள் என்றும், கிராமப்புற நகர்ப்புற மாணவர்கள் என்றும், விடுதி விடுதியல்லாத மாணவர்கள் என்றும் பாடம் மற்றும் வகுப்பு வாரியான மாணவர்கள் என்றும் பிரித்துக் கொண்டு அதன்படி முதன்மைப் பட்டியலில் பதிவு செய்தால் நேர விரயத்தைக் குறைக்கலாம்.

மேலே கூறியவாறு சரியாக முதன்மைப் பட்டியலை அமைத்து விட்டால் ஆய்வு அட்டவணைகளை (Analysis Table) எளிதாகத் தயாரிக்க முடியும். இப்போது சமூக அறிவியலுக்கான புள்ளியியல் சிப்பங்கள் (Statistical Packages for Social Sciences - SPSS) வந்த பிறகு விவரங்களைப் பாகுபடுத்துதலும், பிரித்தலும், அட்டவணைப்படுத்துதலும் மிகவும் எளிதாகிவிட்டது. உதாரணத்திற்கு, நிலங்களின் அளவுப்படி வீடுகளைப் பிரிக்காமலேயே விவரங்களை அட்டவணைப்படுத்தியிருந்தாலும், மேலே கூறப்பட்டுள்ள சிப்பம் எளிதாக நிலத்தின் அடிப்படையில் ஏறு அல்லது இறங்கு வரிசையில் வீடுகளை வரிசைப்படுத்திக் கொடுத்து

விடுகிறது. இது போலவே சாதி அடிப்படையிலோ மதத்தின் அடிப்படையிலோ பிரிக்கச் சொன்னாலும் அந்தச் சிப்பம் பிரித்து ஒரே வகையான சாதி, மதம் சார்ந்த வீடுகள் உள்ள குழுக்களாகப் பிரித்து அட்டவணைப்படுத்தி கொடுத்து விடுகிறது. அதுபோலவே வரைபடங்களையும் தந்து விடுகிறது. இப்படிச் செய்தபின்னர் ஆய்வு அட்டவணைகள் தயாரிப்பது எளிதாகிவிடுகிறது.

மேலும், சேகரிக்கப்பட்ட விவரங்களைத் தேவைக்கேற்ப பிரிக்கவும் மேலே சொன்ன சிப்பம் உதவுகிறது. உதாரணத்திற்கு, புள்ளி விவரங்களை இடத்தின் அடிப்படையிலும் (Spatial or geographical) காலத்தின் அடிப்படையிலும் (temporal or chronological) பண்பினத்தின் அடிப்படையிலும் (qualitative) அளவினத்தின் அடிப்படையிலும் (quantitative) பிரித்து அட்டவணைப் படுத்தலாம்.

பத்தி பத்திகளாகவும் பக்கம் பக்கங்களாகவும் எழுதுகின்ற செய்திகளை அட்டவணைப்படுத்துதல் மூலம் எளிதாக புரியச் செய்துவிடலாம். மேலும், சரியாகப் புள்ளி விவரங்களை அட்டவணைப்படுத்தும்போது, அந்தப் புள்ளி விவரங்கள் சொல்கின்ற உண்மைகள் இன்னும் தெளிவாக விளங்கும். அட்டவணைப்படுத்தப்பட்ட புள்ளி விபரங்களுக்கு இடையேயான உறவுகளும் தெளிவாகும். எனவே புள்ளி விவரங்களை அட்டவணைப்படுத்துதலை ஆய்வின் முதற்படி என்று கூறினால் அது மிகையாகாது.

புள்ளியியலில் ஒரு வகுப்பில் உள்ள 80 மாணவர்கள் பெற்ற மதிப்பெண்கள் கீழேதரப்பட்டுள்ளன(அட்டவணை 1).

அட்டவணை - 1

80 மாணவர்களின் மதிப்பெண்கள்

68	84	75	82	68	90	62	80	76	93
73	79	88	73	60	93	71	59	85	75
61	65	75	87	74	62	95	78	63	72
66	78	82	75	94	77	69	74	68	60
96	78	89	61	75	95	60	79	83	71
79	62	67	97	78	85	76	65	71	75
65	80	73	57	88	78	62	76	53	74
86	67	73	81	72	63	76	75	85	77

குறிப்பு : மேலே கூறப்பட்டுள்ள மதிப்பெண்கள் மொத்த மதிப்பெண்ணான 100க்கு மாணவர்கள் பெற்ற மதிப்பெண்கள்

இந்த 80 எண்களையும் பார்க்கும்போது அந்த எண்களைப் பற்றிய விவரங்கள் அவ்வளவு எளிதாகப் புரிவதில்லை. அவை சொல்லும் செய்திகளும் தெரிவதில்லை. ஏதோ எண்களைக் கொட்டிப் போட்டதுபோல் தோற்றமளிக்கிறது. ஒரு மாத நாள்காட்டியிலும் (Calendar) இப்படித்தான் எண்கள் தோன்றும். ஆனால் ஒரு மாத நாள்காட்டியில் ஏறு வரிசையில் எண்கள் அடுக்கப் பட்டிருக்கும். அந்த எண்கள் நாட்களைக் காட்டுகின்றனவே யொழிய, வேறு எந்த அர்த்தமும் காட்டுவது இல்லை. உதாரணத்திற்கு, ஒரு மாதத்தின் 3ஆம் நாள் வியாழக்கிழமை ஆகவும் 11ஆம் நாள் வெள்ளிக்கிழமை ஆகவும் இருந்தால், 11ஆம் நாள் 3ஆம் நாளைவிடப் பெரியது என்றோ சிறந்தது என்றோ தரத்தில் உயர்ந்தது என்றோ கூறிவிட முடியாது. அந்த 3ம் 11ம் ஒரு வரிசையைக் குறிக்கின்றனவேயொழிய அந்த எண்கள் வேறு அர்த்தம் ஏதும் தரவில்லை. எனவே, நாள்காட்டியில் உள்ள எண்கள் புள்ளிவிவரங்களாகா. ஆனால் அட்டவணை 1இல் உள்ள முதல் இரு எண்கள் 68 மற்றும் 84 ஒரு பொருளைக் குறிக்கின்றன. மதிப்பெண் 68 பெற்ற மாணவரைவிட மதிப்பெண் 84 பெற்ற மாணவர் தேர்வில்

நன்றாகச் செய்திருக்கின்றார் (இரு மதிப்பெண்களையும் கொடுத்த ஆசிரியர்கள் அவர்கள் கடமையினைச் சரியாகச் செய்துள்ளார்கள் என்ற அநுமானத்தில்) என்று சொல்லலாம்.

கொடுக்கப்பட்டுள்ள மதிப்பெண்கள் பற்றிய விவரங்களிலிருந்து மதிப்பெண் 53 பெற்றவர்தான் மிகக் குறைந்த மதிப்பெண் பெற்றவர் என்று கூற இயலும். இவர் அட்டவணை - 1ன்படி ஏழாவது நிரை (Row)யில் ஒன்பதாவது நிரலில் (Column) இருந்தாலும், மதிப்பெண்கள் அடிப்படையில் கடைசியில் இருக்கின்றார். எனவே, இந்த வகுப்பில் உள்ள மாணவர்களில் இந்தப் புள்ளியியல் தேர்வைப் பொறுத்த வரையில், 53 மதிப்பெண்கள் பெற்றவரின் செயல்திறன் மிகக் குறைந்ததாகக் கருதலாம். மிக அதிகமாக 97 மதிப்பெண்கள் பெற்றவர் இந்தத் தேர்வில் மிகச் சிறந்தவராகக் கருதப்படலாம். முன்னால் கொடுக்கப்பட்டுள்ள அட்டவணை 1ல் இந்த மதிப்பெண்களைத் தேடிக்கண்டுபிடிக்க வேண்டியுள்ளது. கொடுக்கப்பட்டுள்ள மதிப்பெண்களை ஏறுவரிசையிலோ இறங்கு வரிசையிலோ அடுக்கியிருந்தால், அந்த அட்டவணையில் தேவையான விபரங்களைத் தேடிக் கண்டுபிடிப்பது மிக எளிதாகியிருக்கும். இதனை எக்ஸெல் (EXCEL) சிப்பம் எளிதாக்கியிருக்கிறது.

அட்டவணை - 2

80 மாணவர்கள் பெற்ற மதிப்பெண்கள் (ஏறு வரிசையில்)

53	57	59	60	60	60	61	61	62	62
62	62	63	63	65	65	65	66	67	67
68	68	68	69	71	71	71	72	72	73
73	73	73	74	74	74	75	75	75	75
75	75	75	76	76	76	76	77	77	78
78	78	78	78	79	79	79	80	81	82
82	83	84	85	85	85	86	87	88	88
88	89	90	93	93	94	95	95	96	97

அட்டவணை 1ஐவிட அட்டவணை 2 பார்ப்பவருக்கு எளிதாகப் பல செய்திகளைத் தர முடியும். எனவே சேகரிக்கப்பட்ட புள்ளி விவரங்களை இவ்வாறாக அடுக்கி வைப்பது நல்லது. முன்னரே கூறியதுபோல இச்செயல்களை இப்போது ச.அ.பு.சி. (SPSS) எளிமைப்படுத்தி விடுகிறது.

அட்டவணை 2ஐ இன்னும் சிறப்பாகச் செய்ய முடியுமா? இன்னும் எளிய வகையில் விபரங்களைப் பெறும்படிச் செய்ய முடியுமா? முடியும். உதாரணத்திற்கு 60 அல்லது 60 மதிப்பெண்களுக்கும்மேல் பெற்ற மாணவர்கள்தான் தேர்வில் தேறியவர்கள் என்றும் 80 அல்லது 80க்கும் அதிகமாக மதிப்பெண்கள் பெற்ற மாணவர்களுக்கு உதவித்தொகை ரூ.2,000 வழங்க வேண்டும் என்றும் இருந்தால், இந்த விபரங்களைக் கொண்டு அட்டவணை அமைத்தால் அட்டவணை இன்னும் சிறப்பாக அமையும் (அட்டவணை 3).

அட்டவணை - 3

தேர்ச்சி பெற்ற, பெறாத மற்றும் உதவித்தொகை பெறும் மாணவர்களின் விபரங்கள்

விளக்கம்	மதிப்பெண்கள்	மாணவர்களின் எண்ணிக்கை
தேறாதவர்கள்	≤ 59	3
தேறியவர்கள்	60 - 79	54
தேறியவர்களில் உதவித்தொகை பெறத் தகுதியானவர்கள்	80 ≤	23
மொத்தம்		**80**

இந்த வகுப்பில் உள்ள 80 மாணவர்களில் 50 பேர் விடுதியில் உள்ளவர்கள் என்றும் 30 பேர் விடுதியில் தங்காதவர்கள் என்றும் பிரிக்கலாம். மாணவர்களை ஆண்கள் பெண்கள் என்றும் கிராமப்புறத்திலிருந்து வருபவர்கள்

நகர்ப்புறத்திலிருந்து வருபவர்கள் என்றும் பிரித்தால் இன்னும் அட்டவணை சிறப்படையும்; நிறையச் செய்திகளையும் தரும். அப்படி வரும்போது அட்டவணைகள், ஒருவழி, இரு வழி, பலவழி அட்டவணைகளாக அமைக்கப்படுகின்றன. உதாரணத்திற்கு அட்டவணை 4ஐப் பார்க்கலாம்.

அட்டவணை - 4
மாணவர்களின் விபரங்கள்

விபரங்கள்	மாணவர்களின் எண்ணிக்கை		மொத்தம்
	விடுதியில் உள்ளவர்	விடுதியில் இல்லாதவர்	
தேறாதவர்கள்	2	1	3
ஆண்கள்	1	1	2
பெண்கள்	1	0	1
தேறியவர்கள்	24	30	54
ஆண்கள்	10	12	22
பெண்கள்	14	18	32
தேறிய மற்றும் உதவித்தொகை பெறத் தகுதியானவர்கள்	11	12	23
ஆண்கள்	5	7	12
பெண்கள்	6	5	11

கொடுக்கப்பட்டுள்ள எடுத்துக்காட்டில் உள்ள மாணவர்களை மதிப்பெண்படியும் பல குழுக்களாகப் பிரிக்கலாம். அப்படிப் பிரிக்கும்போது இரண்டு கேள்விகள் எழலாம். (1) எத்தனை குழுக்களாகப் பிரிப்பது? (2)ஒவ்வொரு பிரிவு இடைவெளியையும் எவ்வளவாக வைப்பது? இந்த மாணவர்களை மறு ஆய்வுக்குப் பயன்படுத்தும் முன்பாக இந்தக் கேள்விகளுக்குப் பதில் கண்டுபிடித்து அதன்படிச் செயல்படுவது நல்லது. இந்த மாணவர்களின் மதிப்பெண்களின் பரவல் இந்த இரு கேள்விகளுக்கும் உள்ள

பதில்களைப் பொறுத்தேதான் அமைகிறது. ஆனால், தற்போது பலரும் இக்கேள்விகளுக்கு முக்கியத்துவம் கொடுக்காமல் அட்டவணை அமைத்து விடுகிறார்கள். அப்படிச் செய்வதால் ஒரேவகையான புள்ளிவிவரத்திற்கு பல வித்தியாசமான பரவல்களும் அதனைத் தொடர்ந்து வித்தியாசமான சராசரிகளும், திட்ட விலக்கங்களும் அமைகின்றன. இது சரியான ஆய்வாக அமையாது. இப்படிச் செய்தால் ஆய்வின் முடிவுகள் பயன்படுத்தக்கூடிய வகையில் அமையாது.

கொடுக்கப்பட்டுள்ள புள்ளி விபரங்களை எத்தனை பிரிவுகளாகப் பிரிக்கலாம் என்பதற்கு ஒரு சூத்திரம் கொடுக்கப்படுகிறது.

பிரிவுகளின் எண்ணிக்கை = $K = 1 + (3.32 \log n)$

n = புள்ளிகளின் எண்ணிக்கை

இப்பொழுது எடுக்கப்பட்டுள்ள உதாரணத்தில் 80 மாணவர்களின் மதிப்பெண்கள் உள்ளன.

எனவே $K = 1 + (3.32 \log 80)$

$= 1 + (3.32 \times 1.90) = 1 + 6.31 = 7.31$

80 மதிப்பெண்களை ஏழு பிரிவுகளாகப் பிரிக்கலாம்.

மாணவர்களின் மதிப்பெண்களுக்கிடையே உள்ள வித்தியாசம் = 97 - 53 = 44. இதை ஏழு பிரிவுகளாகப் பிரித்தால் (44/7 = 6.28) ஒவ்வொரு பிரிவுக்கும் உள்ள இடைவெளியை 6.3ஆகக் கொள்ளலாம். கீழ் எல்லை = 53. மேல் எல்லை = 59.3. இவ்வாறாக மற்ற பிரிவுகளையும் முடிவு செய்யலாம். இப்படிப்பட்ட முடிவினை எடுப்பதால் இந்த புள்ளி விபரத்தினைப் பிரித்து ஆயும் யாவரும் ஒரே மாதிரியான பரவலைப் பெற்று ஆய்வுகளின் முடிவுகள் வேறுபாடின்றி அமையும். இதுபோல, கொடுக்கப்பெற்றுள்ள புள்ளி விபரத்திற்கு சராசரியையும் ($\overline{X}$) , திட்ட விலக்கத்தையும் (σ)

கண்டுபிடித்து $\overline{x} \pm \sigma$ என்று கொண்டு பிரிவுகளின் இடைவெளிகளை திட்டவிலக்கத்தை வைத்து நிர்ணயிக்கலாம். இம்முறையும் வேறுபாட்ற்ற பரவலைத் தரும். மேலே கணக்கிட்டபடி, பிரிவுகளும் அலைவெண்பரவலும் அட்டவணை 5இல் தரப்படுகின்றன.

அட்டவணை - 5
அலைவெண் பரவல்

மதிப்பெண்களின் பிரிவு	மாணவர்களின் எண்ணிக்கை
53.0 - 59.3	3
59.3 - 65.6	14
65.6 - 71.9	10
71.9 - 78.2	27
78.2 - 84.5	9
84.5 - 90.8	10
90.8 - 97.1	7
மொத்தம்	80

கொடுக்கப்பட்டுள்ள மதிப்பெண்கள் முழு எண்களாக இருப்பதாலும், கீழ் மற்றும் மேல் எல்லைகள் ஒரு தசம இலக்கத்தோடு இருப்பதாலும் குழப்பமின்றி அலைவெண்களைக் கண்டுபிடிக்க முடிந்தது. உதாரணத்திற்கு, முதல் பிரிவின் மேல் எல்லையும், இரண்டாம் பிரிவின் கீழ் எல்லையும் 59 ஆக இருந்திருந்தால், ஒரு மாணவனின் மதிப்பெண்ணும் 59ஆக இருந்திருந்தால் அந்த மாணவனை எந்தப் பிரிவில் சேர்ப்பது? முதல் பிரிவிலா? இரண்டாவது பிரிவிலா? இப்படிப் பிரச்சனை வந்தால், கூட ஒரு தசம இலக்கத்தை எடுத்து அந்தப் பிரச்சனையைத் தவிர்த்து விடலாம்.

சில சமயம் முதல் பிரிவின் கீழ் எல்லையும் கடைசிப் பிரிவின் மேல் எல்லையும் குறிப்பிட்டுக் கொடுக்கப்படாமல் முதல் பிரிவின் கீழ் எல்லை ≤ 20 என்றோ, கடைசிப் பிரிவின் மேல் எல்லை 100 ≤ என்றோ குறிப்பிடப்பட்டிருக்கலாம். இப்படிக் கொடுக்கப்பட்டிருந்தால், இடைநிலை, முகடு போன்றவற்றைக் கண்டுபிடிக்க முடியும்; கூட்டுச் சராசரியைக் கண்டுபிடிக்க முடியாது. மேலும் சில சமயம் பிரிவு இடைவெளிகள் ஒவ்வொரு பிரிவுக்கும் வித்தியாசமாக (உ.ம். 10-20, 20-50, 50-100, 100-120, 120-130) இருக்கலாம். அப்படியின்றி எப்பொழுதும் பிரிவு இடைவெளிகள் சமமாக இருத்தல் நன்று.

குவிவு அலைவெண் பரவல்கள்

சில சமயம் ஒரு குறிப்பிட்ட மதிப்பெண்ணைக் கொடுத்து அதற்குக் கீழே அல்லது மேலே எத்தனை மாணவர்கள் மதிப்பெண்கள் பெற்றுள்ளனர் என்று கேட்கலாம். அல்லது, ஒருவர் தினமும் தான் செலவு செய்த பணத்தின் அளவைக் குறித்துக் கொண்டு வரும்போது ஒரு குறிப்பிட்ட நாள் முதல் இன்னொரு குறிப்பிட்ட நாள் வரை எவ்வளவு செலவு செய்துள்ளார் என்று அறிய விரும்பலாம். அல்லது பல நாட்கள் குறித்து வைத்திருந்த செலவு கணக்கைப் பார்த்துவிட்டு இன்றுவரை எவ்வளவு செலவு செய்துள்ளார் என அறிய விரும்பலாம். இப்படி ஒவ்வொரு வாரமும் கூட அறிய விரும்பலாம். இந்த மாதிரி சூழ்நிலைகளில் குவிவு அலைவெண் தயாரிக்கும் முறை வேண்டிய செய்தியை எளிதாகப் பெற உதவும். உதாரணத்திற்கு ஒரு செலவு கணக்கு கொடுக்கப்படுகிறது (அட்டவணை 6).

அட்டவணை - 6
செலவு விபரம்

நாள்	செலவு தொகை (ரூ.)	குவிவு செலவு
1	8	8
2	10	18
3	7	25
4	6	31
5	5	36
6	6	42
7	8	50
8	9	59
9	10	69
10	8	77
11	6	83
12	8	91
13	9	100
14	8	108
15	10	118
16	12	130
17	18	148
18	10	158
19	9	167
20	8	175
21	2	177
22	7	184

அட்டவணை 6இல் கொடுக்கப்பட்ட செலவு விபரத்திலிருந்து முதல் 5 நாட்களில் செலவு செய்யப்பட்ட தொகை எவ்வளவு என்றால் கூட்டிப் பார்க்காமலே உடனேயே ரூ.36 என்று கூறிவிடலாம். அல்லது முதல் 10 நாட்களில் செலவு செய்யப்பட்ட தொகை எவ்வளவு என்றால் மறுபடியும் கூட்டிப் பார்க்காமல் உடனே ரூ.77 என்று சொல்லி விடலாம். இவ்வாறாக, குவிவு செலவு கணக்கையும் கூடவே எழுதி வைத்து விட்டால் பல தடவைகள் கூட்டிச் சொல்வதைத் தவிர்த்துவிட்டு பார்த்த மாத்திரமே உடனே கேள்விக்குப் பதிலைக் கூறிவிடலாம். இப்போது குவிவு அலைவெண் பரவலுக்கு அட்டவணை 5ஐ எடுத்துக் கீழே கொடுக்கப்பட்டுள்ள அட்டவணை 7ஐ உருவாக்கலாம்.

அட்டவணை - 7
குவிவு அலைவெண் பரவல்

மதிப்பெண்கள் பிரிவு	மாணவர்களின் எண்ணிக்கை	கீழினக் குவிவு அலைவெண்
53.0 - 59.3	3	3
59.3 - 65.6	14	17
65.6 - 71.9	10	27
71.9 - 78.2	27	54
78.2 - 84.5	9	63
84.5 - 90.8	10	73
90.8 - 97.1	7	80

இந்த அட்டவணையிலிருந்து 71.9க்கும் கீழ் எத்தனை மாணவர்கள் மதிப்பெண்கள் பெற்றுள்ளனர் என்று கேட்டால், பார்த்த உடனேயே 27 பேர் என்று கூறிவிடலாம். அல்லது 84.5க்கும் கீழ் எத்தனை மாணவர்கள் மதிப்பெண்கள் பெற்றுள்ளனர் எனக் கேட்டாலும் உடனே 63 மாணவர்கள் என்று கூறிவிடலாம். இந்த அட்டவணை ஒரு குறிப்பிட்ட

மதிப்பெண்ணுக்குக் கீழே பெற்றுள்ள மாணவர்களின் எண்ணிக்கையை உடனே கூற உதவுவதால் இதைக் கீழினக் குவிவு அலைவெண் பரவல் எனலாம். இதுபோல மேலினக் குவிவு அலைவெண் பரவலும் தயாரிக்கலாம். இதை அட்டவணை 8ல் காணலாம்.

அட்டவணை - 8

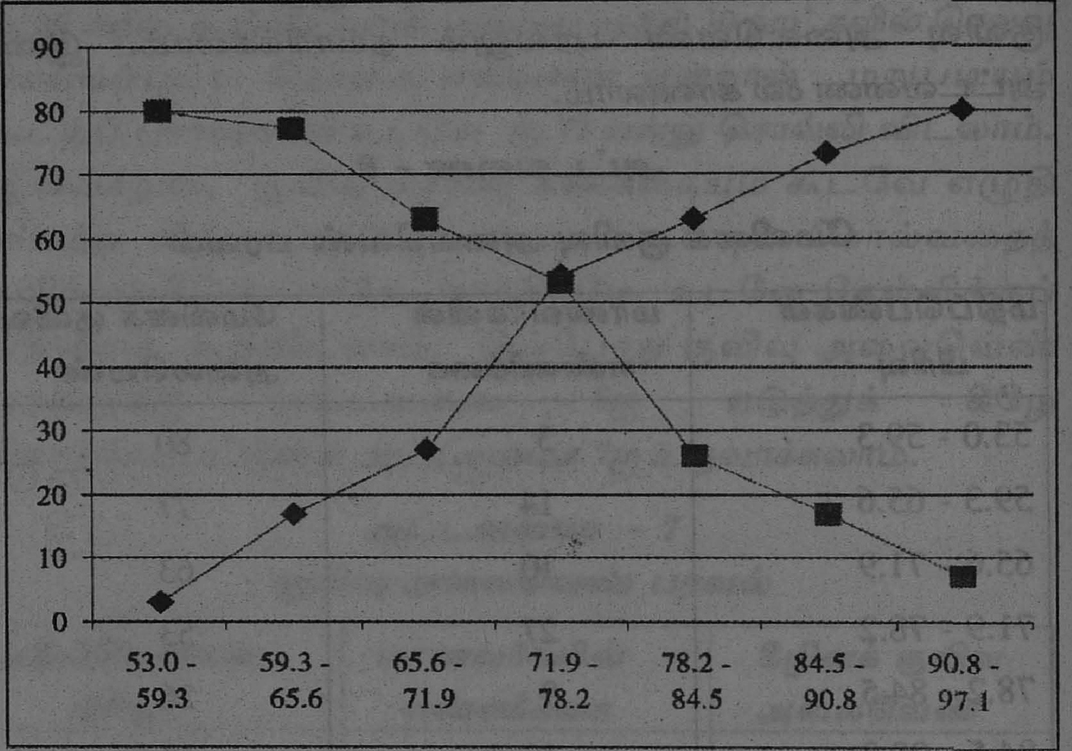
மேலினக் குவிவு அலைவெண் பரவல்

மதிப்பெண்கள் பிரிவு	மாணவர்களின் எண்ணிக்கை	மேலினக் குவிவு அலைவெண்
53.0 - 59.3	3	80
59.3 - 65.6	14	77
65.6 - 71.9	10	63
71.9 - 78.2	27	53
78.2 - 84.5	9	26
84.5 - 90.8	10	17
90.8 - 97.1	7	7

இந்த அட்டவணை ஒரு குறிப்பிட்ட மதிப்பெண்ணுக்கும் மேல் எத்தனை மாணவர்கள் மதிப்பெண் பெற்றுள்ளனர் என்று கேட்டால் எளிதாகப் பதில் சொல்லும் வண்ணம் அமைந்துள்ளது. உதாரணத்திற்கு, 65.6 மதிப்பெண்களுக்கும் மேல் எத்தனை மாணவர்கள் மதிப்பெண்கள் பெற்றுள்ளனர் என்று கேட்டால், பார்த்த மாத்திரமே 63 மாணவர்கள் எனக் கூறிவிடலாம்.

மேலின மற்றும் கீழின அலைவெண்களைப் பயன்படுத்தி வரைபடங்களும், விளக்கப்படங்களும் வளைகோடுகளும் வரையலாம். அவ்விரண்டு வளைகோடுகளும் சந்திக்கும் இடத்திலிருந்து அந்தப் பரவலின் இடைநிலையையும் பெறலாம்.

வரைபடம் - 1
அலைவெண்கள்



அலைவெண் வளைகோடுகள் அல்லது நேர்கோடுகள் சில சமயங்களில் பொருத்தமில்லாத இடங்களில் பயன்படுத்தப்படுகின்றன. உதாரணத்திற்கு, தொடரில்லா மாறிக்கு இதைப் பயன்படுத்துவது சரியல்ல. தொடர் மாறிக்கு (Continuous variable) இதைப் பயன்படுத்துவது பொருத்தமாகும். ஒரு வகுப்பு மாணவர்கள் எடுத்த மதிப்பெண்கள் 0விலிருந்து 100 வரை இருந்தால் ஒவ்வொரு மதிப்பெண்ணும் எத்தனை மாணவர்களால் எடுக்கப்பட்டுள்ளது என்று பார்க்கலாம். அல்லது 0விலிருந்து 10 வரை; 10 முதல் 20 வரை என்று பார்த்து ஒவ்வொரு இடைவெளியிலும் எத்தனை மாணவர்கள் உள்ளனர் என்று பார்க்கலாம். அல்லது, ஒரு நாடு ஒவ்வொரு ஆண்டும் எந்த அளவுக்கு ஏற்றுமதி செய்துள்ளது எனக் காலம்சார் தொடராக (Time Series Data and Continuous

Variable) ஆய்வு செய்யலாம். இந்தச் சூழ்நிலைகளில் பட்டை வரைபடம் வரைவதும், ஒவ்வொரு பட்டையின் மையப் புள்ளியையும் இணைத்து வரைகோடு (அலைவெண் வளைகோடு) வரைவதும் பொருத்தமாகும்.

அவ்வாறில்லாமல், பல நாடுகள் ஓர் ஆண்டு ஏற்றுமதி செய்ததைப் பட்டை வரைபடமாக வரைந்து அப்பட்டைகளின் மையப்புள்ளிகளை இணைத்து வளைகோடுகள் வரைவது பொருந்தாததாகும். நான்கு வகுப்பில் உள்ள மாணவர்களின் மதிப்பெண்களை (உதாரணத்திற்கு 30, 50, 40, 20 என்பதை) பட்டை வரைபடமாக வரையலாம். அந்த மதிப்பெண்களை ஒரு கோட்டின் மூலம் இணைப்பது பொருத்தமில்லாததாகும். இந்தியாவின் ஏற்றுமதியை 1950 முதல் 1980 வரை; 1980 முதல் 1990 வரை; 1990 முதல் 2009 வரை பிரித்து ஒவ்வொரு காலகட்டத்திற்கும் சராசரி பார்த்து அவற்றைச் செவ்வக வரைபடமாக (ஆண்டுகளை அடியாகவும், சராசரியை உயரமாகவும் கொண்டு) வரையலாம். ஆனால், அவற்றை வளைகோடு போட்டு ஒவ்வொரு ஆண்டும் ஏற்றுமதி எப்படியிருந்தது என்று காட்டவேண்டுமென நினைத்தால், ஒவ்வொரு ஆண்டுக்கும் உள்ள உண்மையான புள்ளிகளை வரைபடத்தில் காட்டி, அவற்றை இணைத்து வளைகோடு வரைவது சரியாகும். அல்லது முடிந்தால் சாதாரண மிகக் குறைந்த வர்க்க (Ordinary least square = OLS) முறையைப் பயன்படுத்தி நேர்கோடும் வரையலாம். அந்த நேர்கோட்டில் இருந்து உண்மையான புள்ளிகள் எந்தளவுக்குத் தள்ளி உள்ளன? அதன் பொருள் என்ன? ஏன் அப்படித் தள்ளி உள்ளன? போன்ற கேள்விகளை எழுப்பி அதற்கு பதில் கண்டுபிடிப்பது ஒரு நல்ல ஆய்வின் குறிக்கோளாக இருக்கும்.

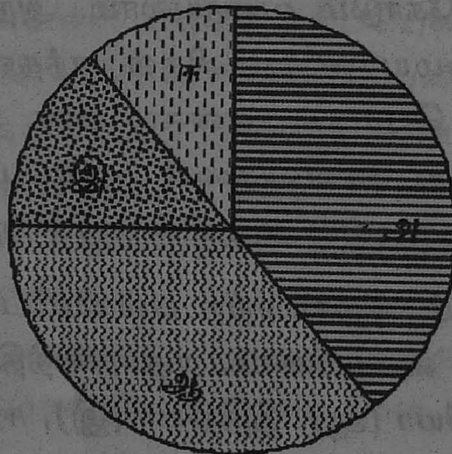
சில சமயம் தொடரற்ற புள்ளிவிபரங்களை வட்ட வரைபடமாகவும் வரையலாம். உதாரணத்திற்கு, அமெரிக்கா (அ), ஆஸ்திரேலியா (ஆ), இந்தியா (இ), ஈரான் (ஈ) போன்ற

நாடுகளின் இறக்குமதிக்கான புள்ளிவிபரங்களை பத்திகளில் எழுதிக்காட்டுவதைவிட, ஒரு வட்ட வரைபடமாகக் காட்டினால் எளிதில் புரிந்து கொள்ள முடியும். அதற்கு நான்கு நாடுகளின் இறக்குமதியையும் ஒரு வட்டத்திற்கான மொத்த கோணமாகிய 360க்கு மாற்றி தனித்தனியாக ஒவ்வொரு நாட்டிற்கும் கோணத்தைக் கணக்கிட்டு வட்டத்திற்குள் வெவ்வேறு வண்ணத்தால் காட்டலாம். உதாரணத்திற்கு, அந்த நாடுகளின் இறக்குமதி முறையே ரூ.1400 கோடி, ரூ.1300 கோடி, ரூ.500 கோடி, ரூ.400 கோடி என்று வைத்துக் கொண்டால் அவற்றை 140, 130, 50, 40 என்று கோணங்களாக மாற்றி வட்ட விளக்கப் படங்களாகக் காட்டலாம்.

விளக்கப்படங்கள்

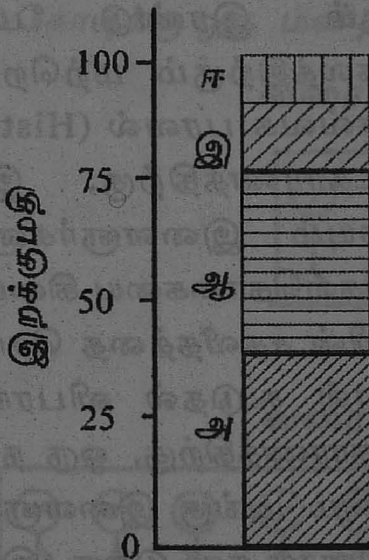
ஆனால், சில நாடுகளின் படிப்பறிவு பெற்றோர் விகிதமோ (literacy ratio) அல்லது ஆண்-பெண் விகிதமோ (sex ratio) கொடுத்திருந்தால் அவற்றை ஒரு வட்டவிளக்கப் படத்தில் காட்டுவது சரியாகாது. அவற்றை ஒவ்வொரு நாட்டுக்கும் தனி பட்டை வரைந்து பட்டை வரைபடத்திலோ, சதவீதப்பட்டை வரைபடத்திலோ காட்டலாம். வட்டவிளக்கப் படத்தில் வரையப்படும் மாறி கூட்டப்படக்கூடியதாக இருந்து அதன் மொத்தம் 360க்கு மாற்றக்கூடியதாகவும் இருக்க வேண்டும்.

வரைபடம் - 2



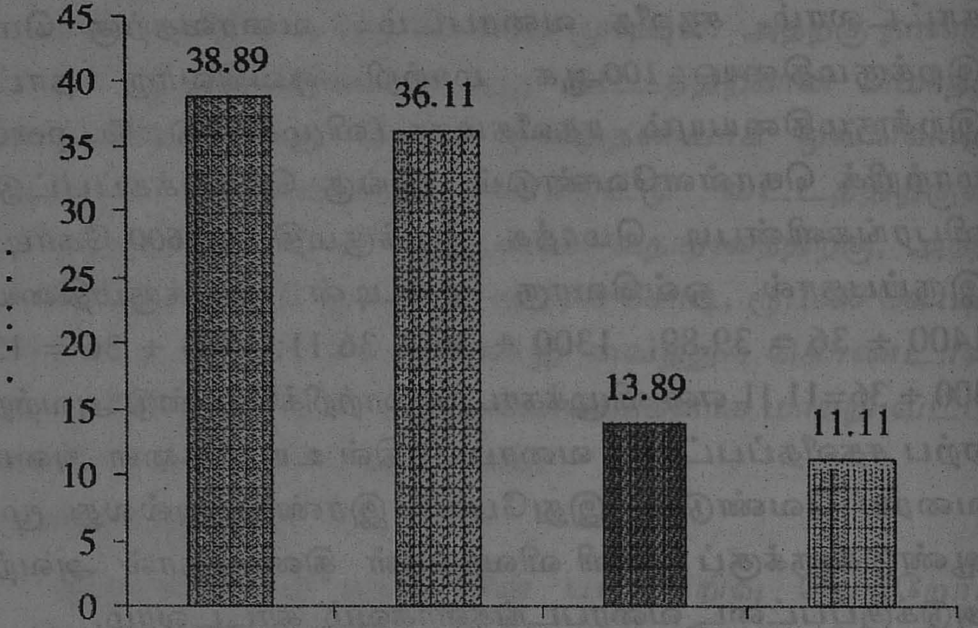
மேற்கூறப்பட்ட செய்தியைப் பட்டை வரைபடமாகவோ, சதவீதப்பட்டை வரைபடமாகவோ காட்டலாம். சதவீத வரைபடம் வரைவதற்கு மொத்த இறக்குமதியை 100ஆக மாற்றி ஒவ்வொரு நாட்டின் இறக்குமதியையும் சதவீதமாக (விழுக்காடு, %, percent) மாற்றிக் கொள்ளவேண்டும். இங்கு கொடுக்கப்பட்டுள்ள விபரங்களின்படி மொத்த இறக்குமதி ரூ.3600 கோடியாக இருப்பதால் ஒவ்வொரு நாட்டின் இறக்குமதியையும் 1400 ÷ 36 = 39.89; 1300 ÷ 36 = 36.11; 500 ÷ 36 = 13.89; 400 ÷ 36=11.11 என விழுக்காடாக மாற்றிக்கொண்டு அவற்றிற்கு ஏற்ப சதவீதப்பட்டை வரைபடத்தில் உயரங்களை அமைத்து வரைய வேண்டும். இதுபோல இரண்டு அல்லது மூன்று ஆண்டுகளுக்குப் புள்ளி விவரங்கள் கிடைத்தால் அவற்றை அடுக்குப்பட்டை வரைபடங்களாகவும் காட்டலாம்.

வரைபடம் - 3



சதவீதப்பட்டை வரைபடம்

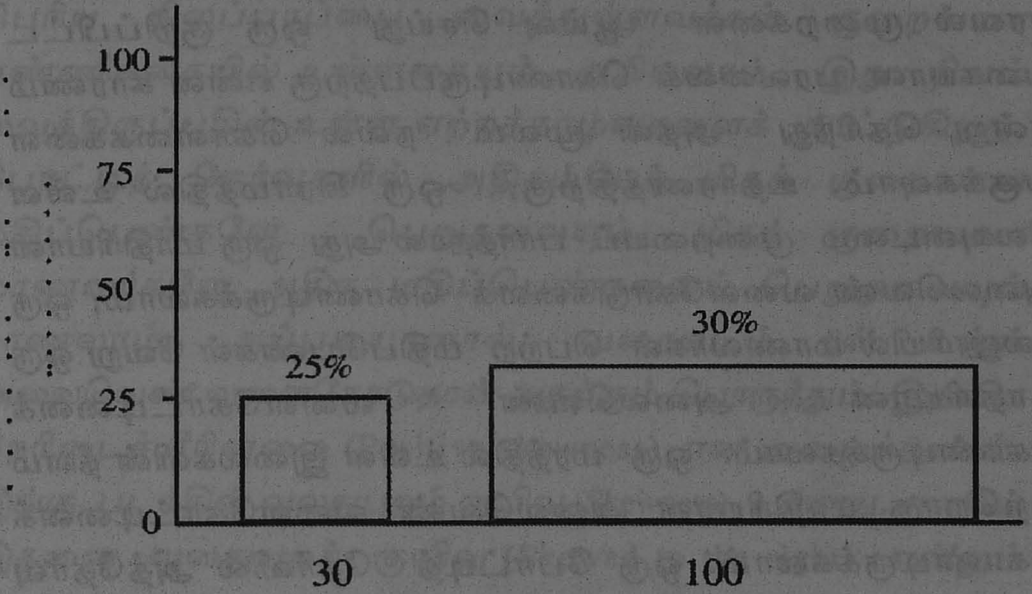
வரைபடம் - 4



அடுக்குப்பட்டை வரைபடம்

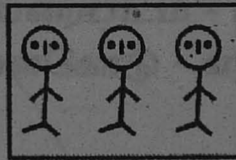
ஒவ்வொரு நாட்டிற்கும் இரண்டு மாறிகள் கொடுத்திருந்தால் ஒன்றை அகலத்திற்கும் மற்றொன்றை உயரத்திற்கும் பயன்படுத்தி செவ்வக பரவல் (Histogram) வரைபடம் வரையலாம். உதாரணத்திற்கு, இரண்டு நாடுகளுடைய மக்கள்தொகையும் இளைஞர்களுடைய சதவீதமும் கொடுத்திருந்தால் மக்கள்தொகையை கிடைமட்ட அச்சிலும் (x axis) இளைஞர்களின் சதவீதத்தை செங்குத்து (y axis) அச்சிலும் கொடுத்தால் கூடுதல் விபரங்களை வரைபடத்திலிருந்து பெறலாம். உதாரணத்திற்கு, ஒரு நாட்டின் (அ) மக்கள்தொகை 30 கோடியாகவும் அங்கு இளைஞர்கள் 25 சதவீதம் உள்ளனர் என்றும் இன்னொரு நாட்டுக்கு (இ) இந்த விபரங்கள் முறையே 100 கோடியாகவும் 30 விழுக்காடாகவும் உள்ளனர் என்றும் கொண்டால் அவை வரைபடம் 5இல் உள்ளது போல் தோற்றமளிக்கும்.

வரைபடம் - 5



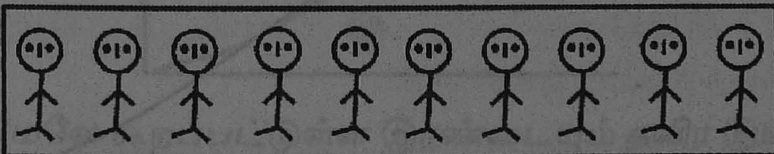
மக்கள்தொகையினைக் காட்டச் சித்திர விளக்கப் படங்களையும் கையாளலாம். மேலே கொடுக்கப்பட்டுள்ள மக்கள்தொகை தொடர்பான விபரங்களைச் சித்திர விளக்கப் படங்கள் மூலம் கீழே கொடுக்கப்பட்டுள்ளவாறு காட்டலாம். இதில் 10 கோடிக்கு ஒரு மனிதத் தலையைப் பயன்படுத்தலாம்.

வரைபடம் - 6



அ

வரைபடம் - 7



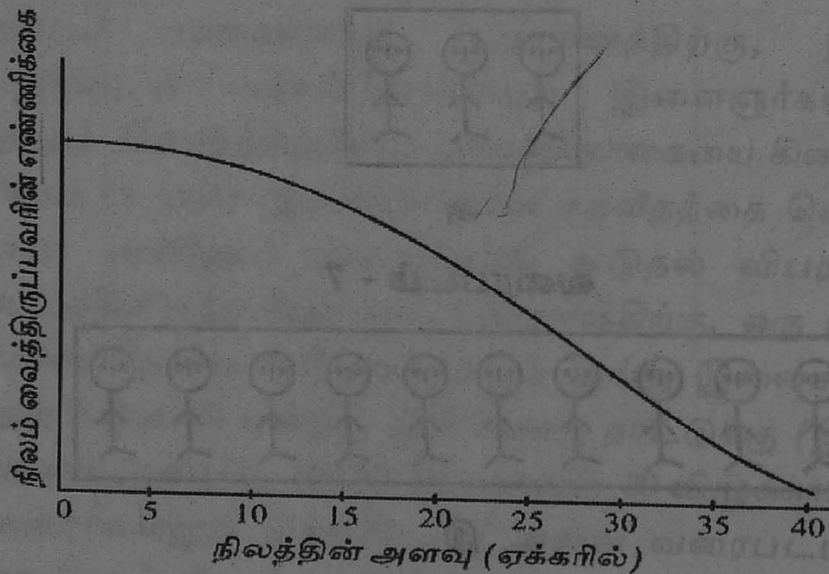
இ

அலைவெண் வளைகோடுகளின் வகைகள்

அலைவெண் வளைகோடுகளின் தன்மைகளை வைத்து பரவல் முறைகளை ஆய்வு செய்து ஒரு குறிப்பிட்ட வகையான பரவலைக் கொண்டிருப்பதற்கு என்ன காரணம் என்று தெரிந்து அதன் மூலம் நல்ல கொள்கைகளை வகுக்கலாம். உதாரணத்திற்கு, ஒரு கிராமத்தில் உள்ள நிலவுடைமை முறையைப் பார்த்தால் அது ஒரு மாதிரியான அலைவெண் வளைகோடுகளைக் கொண்டிருக்கலாம்; ஒரு கல்லூரியில் மாணவர்கள் பெற்ற மதிப்பெண்கள் வேறு ஒரு மாதிரியான அலைவெண் வளைகோட்டினைக் கொண்டிருக்கலாம். ஒரு மரத்தில் உள்ள இலைகளின் நீளம் மற்றொரு மாதிரியான அலைவெண் வளைகோட்டினைக் கொண்டிருக்கலாம். ஒரு போட்டித் தேர்வில் அத்தேர்வு எழுதியவர்கள் பெற்ற மதிப்பெண்கள் இன்னுமொரு மாதிரியான அலைவெண் வளைகோட்டினைக் கொண்டிருக்கலாம். இவ்வாறு வெவ்வேறு வகையான வளைகோடுகளைக் கொண்டிருப்பதற்கான காரணத்தை ஆய்வு செய்து பல உண்மைகளைக் கண்டுபிடிக்க முடியும்.

வரைபடம் - 8

அலைவெண் வளைகோடு - நேரிடைச் சீரின்மை

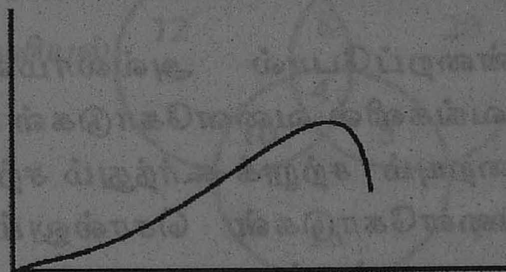


வரைபடம் 8இன் மூலம் சிறிய அளவு நிலம் வைத்திருப்பவர்கள் அதிக எண்ணிக்கையில் இருப்பதையும் பெரிய நிலப்பரப்பை வைத்துள்ளவர்கள் குறைவான எண்ணிக்கையில் உள்ளதையும் அறியலாம். இது நிலம் வைத்திருப்பதில் உள்ள ஏற்றத்தாழ்வுகளைக் காட்டுகிறது. போட்டித் தேர்வுகளில் அதிகம்பேர் மிகக் குறைவான மதிப்பெண்களே பெறுதலையும், மிகக் குறைவான மாணவர்களே அதிக மதிப்பெண்களைப் பெறுதலையும் காணலாம். அப்படியானால், வரைபடம் 8ல் உள்ள அலைவெண் வளைகோடுதான் அதற்கும் பொருந்தும். இதனை நேரிடைச் சீரின்மை (Positive skewness) என அழைக்கலாம். இங்கு படத்தின் வலதுபுறம் குவிவு (Skewness) உள்ளது. எனவே இதனை வலதுபுறக் குவிவு (Skewed to the right) என்றும் அழைக்கலாம்.

ஆனால் கல்லூரிகளிலும் பல பல்கலைக்கழகங்களிலும் அதிகமான மாணவர்கள் அதிக மதிப்பெண்கள் பெறுவதையும் மிகக் குறைவான மாணவர்களே குறைவான மதிப்பெண்கள் பெறுவதையும் காணலாம். இங்கு எதிரிடையான சீரின்மை (negative skewness) காணப்படுகிறது. குவிவு இடதுபுறம் (Skewed to the left) இருக்கும். இதனை வரைபடம் 9ல் காணலாம்.

வரைபடம் - 9

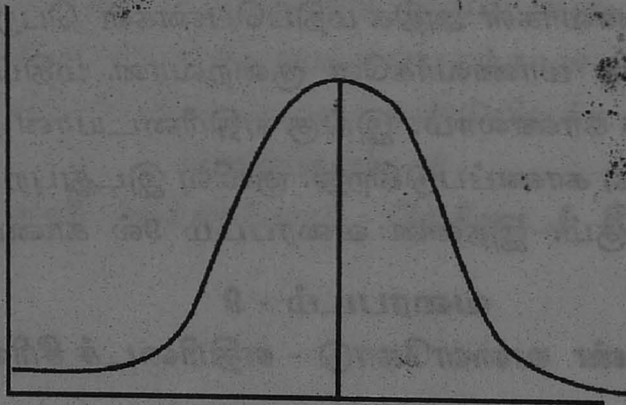
அலைவெண் வளைகோடு - எதிரிடைச் சீரின்மை



மேலே கூறப்பட்டுள்ள இரண்டையும் தவிர இயல்பான பரவல்களும் (Normal distribution) காணலாம். உதாரணத்திற்கு,

ஒரு மரத்தில் உள்ள இலைகளின் நீளத்தைக் கூறலாம். மிகக் குட்டையான இலைகளும் மிக நெட்டையான இலைகளும் எண்ணிக்கையில் குறைவாகவும் ஓரளவுக்குச் சராசரி நீளம் உள்ள இலைகள் எண்ணிக்கையில் அதிகமாகவும் இருக்கலாம். அதுபோல ஒரு குறிப்பிட்ட இடத்தில் உள்ள தட்பவெட்ப நிலைகளைக் கூடச் சொல்லலாம். ஓர் இடத்தில் ஒரு நாளில் காலையில் இருந்து வெப்பம் கொஞ்சம் கொஞ்சமாகக் கூடி, உச்சத்தை அடைந்த பிறகு கொஞ்சம் கொஞ்சமாகக் குறைய ஆரம்பிக்கும். இதனை சமச்சீர் வளைகோடு (Symmetrical, bell-shaped or normal) எனலாம். இது வரைபடம் 10இல் உள்ளது போல் இருக்கும். இதனை இயல்நிலைப் பரவல் என்றும் சொல்வதுண்டு.

வரைபடம் - 10
அலைவெண் வளைகோடு - சமச்சீர் வளைகோடு



இந்த மூன்றைப்போல் அல்லாமல் இன்னும் எத்தனையோ வடிவங்களில் வளைகோடுகள் இருக்கலாம். அவை ஒவ்வொன்றையும் சற்றுக் கூர்ந்தும் சற்று ஆய்ந்தும் பார்த்தால் அவ்வளைகோடுகள் சொல்லும் செய்திகள் சிந்தனையைத் தூண்ட வைக்கும்.

ஒரு பாடத்தில் பெற்ற மதிப்பெண்களை மேலே கூறியவாறு காட்டமுடியும். ஒன்றுக்கும் மேற்பட்ட

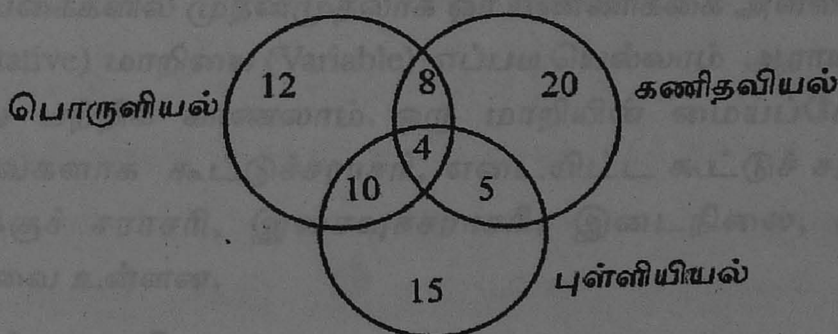
பாடங்களில் மாணவர்கள் தேர்ச்சி பெற்றுள்ள விபரத்தைக் காட்ட யூலர் (Euler) அல்லது வென் (Venn) வரைபடங்களைப் பயன்படுத்தலாம். உதாரணத்திற்கு, அட்டவணை 9ஐப் பார்க்கலாம்.

அட்டவணை - 9
மாணவர்கள் தேர்ச்சி பெற்ற விபரங்கள்

பாடங்கள்	தேர்ச்சி பெற்ற மாணவர்களின் எண்ணிக்கை
பொருளியலில் மட்டும்	12
புள்ளியியலில் மட்டும்	15
கணிதவியலில் மட்டும்	20
புள்ளியியலிலும் பொருளியலிலும்	10
புள்ளியியலிலும் கணிதவியலிலும்	5
பொருளியலிலும் கணிதவியலிலும்	8
மூன்று பாடங்களிலும்	4

இந்த விபரங்களை ஒரு வென் வரைபடம் (11) மூலம் எளிதாகப் புரியும்படி காட்டலாம்.

வரைபடம் - 11



இந்த அட்டவணையில் இன்னும் சில பாடங்களைப் பற்றிய விபரங்களைச் சேர்த்தாலும், அவற்றையும் வென்

வரைபடம் மூலம் தெளிவாகக் காட்டமுடியும். இரண்டு பாடங்கள் மட்டுமே இருந்தால் (உதாரணத்திற்கு பொருளியலையும் புள்ளியியலையும் எடுத்துக்கொள்வோம்) அவற்றை அட்டவணை 10இல் காட்டியதுபோலக் காட்டலாம்.

அட்டவணை - 10

பாடம்வாரியாகத் தேறியவர்களின் எண்ணிக்கை

தேறியவர்களின் எண்ணிக்கை	பொருளியல்	புள்ளியியல்
பொருளியல்	12	10
புள்ளியியல்	10	15

மூன்று அல்லது நான்கு பாடங்கள் இருந்தால் வென் (Venn) வரைபடமே சிறந்தது ஆகும்.

௸

2. சராசரி, இடைநிலை, முகடு மற்றும் பிற மையப்போக்கு அளவீடுகள் (MEASURES OF CENTRAL TENDENCY)

இதுவரையிலும் புள்ளி விபரங்களைச் சேகரிக்கும் முன்பு என்ன முயற்சிகள் மேற்கொள்ளவேண்டும், எப்படிச் சேகரிக்க வேண்டும், சேகரிக்கும்போது கடைப்பிடிக்க வேண்டிய முறைகள் யாவை, சேகரிக்கப்பட்ட விபரங்களை எப்படி ஒழுங்குபடுத்துவது, அவற்றை எப்படி படிப்பவருக்கு எளிதாக்கிக் காட்டுவது போன்ற செய்திகள் விவரிக்கப்பட்டன. சேகரிக்கப்பட்ட விபரங்களை எளிதாகப் புரியும்படி படைப்பதே இதுவரையிலான நோக்கம்; ஆனால், ஆய்வில் இது ஒரு முதல்படியே; அடுத்த நடவடிக்கை என்னவெனில் அந்தப் புள்ளி விவரங்கள் என்ன சொல்ல வருகின்றன என்பதைப் பார்ப்பதாகும். முன்பே கூறியபடி பலவகையான புள்ளிவிபரங்கள் (முதல் மற்றும் இரண்டாம்நிலை விவரங்கள் தவிர) இருக்கின்றன. அவை ஒவ்வொன்றையும் ஆராய்ந்து அவை தருகின்ற செய்திகளைச் சரியாகப் புரிந்து மற்றவர்களுக்குச் சொல்வது ஆய்வாளரின் நோக்கம் ஆகும். இந்த வகைகளில் முதன்முதலாக ஓர் எண்ணிக்கை அளவிலான (Quantitative) மாறியை (Variable) எப்படியெல்லாம் ஆராயலாம் என்பது பற்றிக் காணலாம். ஒரு மாறியின் மையப்போக்கு அளவைகளாக கூட்டுச்சராசரி, எடையிட்ட கூட்டுச் சராசரி, பெருக்குச் சராசரி, இசைவுச்சராசரி, இடைநிலை, முகடு ஆகியவை உள்ளன.

கூட்டுச் சராசரி

ஒரு மாதிரியிலிருந்து கிடைத்த விபரங்களை X (ஓர் ஆங்கில எழுத்து) என்று சொல்லலாம். இந்த X என்பது

மதிப்பெண்ணாகவோ, உயரமாகவோ, எடையாகவோ, வருமானமாகவோ, செலவாகவோ அல்லது இவை போன்ற ஏதாவது ஒன்றாகவோ இருக்கலாம். எடுக்கப்பட்ட மாதிரியில் (Sample) *50* உறுப்புக்கள் (elements) இருந்தால் அவற்றை *50* புள்ளிகளாகக் (Observation) குறிக்கலாம். இதில் மாதிரி (Sample) ஒன்று; புள்ளிகள் *50* (Number of observations = 50). இதை *50* மாதிரிகள் (50 samples) என்று சொல்வது சரியல்ல. இப்படி இருக்கும்போது X_i என்பது முதல் உறுப்பாகவோ, *10*ஆவது உறுப்பாகவோ *50*ஆவது உறுப்பாகவோ அல்லது வேறு எதுவாகவோ இருக்கலாம். i = 10 என்றால் பத்தாவது உறுப்பின் மதிப்பு என்று பெயர். உதாரணத்திற்கு எடுக்கப்பட்ட மாதிரி *50* மாணவர்களைக் கொண்டது என்றால் i ன் மதிப்பு *1* முதல் *50* வரை இருக்கும். ΣX_i (Sigma X_i) என்று எழுதினால், X ஐ மதிப்பெண் என்று கொண்டால், *50* மாணவர்களின் மதிப்பெண்களின் கூடுதல் அல்லது மொத்தம் என்று பொருள். Σ (Sigma) என்பது கிரேக்கத்தில் உள்ள எழுத்துக்களில் ஒன்றான சிக்மாவின் பெரிய எழுத்து (Capital letter). இதுபோல வேறு ஒரு மாறியை Y என்றும் இன்னும் ஒரு மாறியிருந்தால் Z என்றும் கொள்வது வழக்கம். ΣX_i ஐ $\sum_{i=1}^{n} X$ என்றும் ΣX என்று எளிமையாகவும், சுருக்கமாகவும் கூறலாம்.

எனவே $\sum_{i=1}^{n} X = X_1 + X_2 + X_3 + \ldots + X_n$. இதனுடைய கூட்டுச் சராசரி ($\bar{X}$) காண ΣX ஐ அந்தத் தொடரில் உள்ள எண்களின் எண்ணிக்கையால் (n or number of observations) வகுக்க வேண்டும். எனவே $\bar{X} = \frac{\Sigma X}{n}$. இந்த $\bar{X}$க்காக வரும் எண் அந்தத் தொடரில் உள்ள ஓர் எண்ணாகவும் இருக்கலாம்; இல்லாமலும் இருக்கலாம். ஆனால் பொதுவாக அந்தத் தொடரில் இருக்கும் ஓர் எண்ணாக இருக்க வேண்டும்

என்ற கட்டாயம் இல்லை. உதாரணத்திற்கு $X_i = 1, 2, 4, 5$ என்றால் இதன் கூட்டுச் சராசரி $(\bar{X}) = \left[\frac{1+2+4+5}{4} = \frac{12}{4}\right] = 3.$

கூட்டுச் சராசரி (Arithmetic Mean = AM) தவிர இன்னும் சில மையப்போக்கு அளவைகளும் உள்ளன. அவை எடையிட்ட அல்லது நிறையிட்ட கூட்டுச் சராசரி (Weighted Average), பெருக்குச் சராசரி (Geometric Mean), இசைவுச் சராசரி (Harmonic Mean), இடைநிலை, முகடு ஆகியவை. இவையனைத்தும் சராசரிகள் (Averages) என்று அழைக்கப்படுகின்றன. ஆனாலும், இவைகளில் ஒன்றுக்குப் பதிலாக இன்னொன்றைப் பயன்படுத்துவது சரியாக இருக்காது. கிடைத்திருக்கும் புள்ளி விவரங்களின் தன்மைகள் எந்தச் சராசரியைப் பயன்படுத்த வேண்டும் என்று நிர்ணயிக்கின்றன. மேலும், ஆய்வின் நோக்கத்தைப் பொறுத்தும் அது அமையும்.

கிடைத்துள்ள விவரங்களின் தன்மையைப் பொறுத்து கூட்டுச் சராசரி காணும் முறையும் மாறுபடும். உதாரணத்திற்கு, அட்டவணை 2இல் கொடுக்கப்பட்டுள்ள 80 மாணவர்களின் மதிப்பெண்களையும் கூட்டி அதை 80ஆல் வகுத்து கூட்டுச் சராசரி காணலாம். அல்லது ஒவ்வொரு எண்ணும் (X) எத்தனை முறை வருகிறது (frequency = f) என அறிந்து அந்தந்த எண்களை அவற்றிற்குரிய அலைவெண்களால் (f) பெருக்கி (fx) பின்னர் பெருக்கி வந்ததைக் கூட்டி (Σfx) அந்த தொகுதியில் உள்ள எண்களின் எண்ணிக்கையால் (இங்கு n = 80) வகுத்து $\left(\frac{\Sigma fx}{\Sigma f}\right)$ கூட்டுச்சராசரி கணக்கிடலாம். இந்த இருமுறைகளிலும் வருகின்ற கூட்டுச் சராசரிகள் மிகத் துல்லியமாக சமமாக வரவேண்டும். அப்படியில்லையெனில் அங்கு ஏதோ தவறு நடந்துள்ளது என்பது பொருள்.

அட்டவணை 2லிருந்து பெறப்பட்ட அட்டவணை 5இல் உள்ள விபரங்களுக்கும் கூட்டுச் சராசரி கணக்கிடலாம். ஆனால், இந்த மாதிரி கண்டுபிடிக்கப்பட்ட கூட்டுச் சராசரி முதல் இரண்டு முறைகளால் கண்டுபிடித்த அளவுக்குத் துல்லியமாகச் சமமாக வர வாய்ப்பு மிகவும் குறைவு. ஏனெனில், இங்கு புள்ளி விபரங்களை அட்டவணைப் படுத்தும்போது கீழே கொடுக்கப்படும் சில அனுமானங்கள் மேற்கொள்ளப்பட்டுள்ளன. அவை சரியாக இருக்க வேண்டுமென்ற கட்டாயமில்லை. எனவே இந்த மூன்றாவது முறையால் கண்டுபிடித்த கூட்டுச் சராசரி முதல் இரண்டு முறைகளில் கண்டுபிடித்தவற்றை விடச் சற்று வித்தியாசமாக இருக்கலாம்.

அட்டவணை 5இல் உள்ள விபரங்களுக்கு கூட்டுச் சராசரி கணக்கிடும்பொழுது ஒவ்வொரு மதிப்பெண் பிரிவிற்கும் மைய மதிப்பு (Mid-value) கண்டுபிடித்து அதை அதனோடு தொடர்புடைய அலைவெண்ணுடன் பெருக்குகிறோம். இப்படிச் செய்யும்போது, மைய மதிப்பு என்பது அந்த மதிப்பெண் பிரிவுக்குப் பதிலாக எடுத்துக் கொள்ளப்படுகிறது. இது சரியாக இருக்க வேண்டுமென்ற அவசியம் இல்லை. உதாரணத்திற்கு, முதல் மதிப்பெண் பிரிவாகிய 53.0 - 59.3இல் உள்ள 3 எண்களுமே 53 ஆகவோ அல்லது 59 ஆகவோ கூட இருக்கலாம். அதனுடைய மைய மதிப்பை எடுத்து அதன் அலைவெண்ணுடன் பெருக்கி வருகின்ற எண் உண்மையான ($53 \times 3 = 159$) எண்ணைவிடக் கூடவோ அல்லது ($59 \times 3 = 177$) குறைவாகவோ இருக்கலாம். அப்படி இருக்கும் நிலையில், மைய மதிப்பை அலைவெண்ணோடு பெருக்கிக் கண்டுபிடிக்கும் கூட்டுச் சராசரி அந்த எண்களின் உண்மையான கூட்டுச் சராசரியை விட வித்தியாசமாக இருக்கலாம்.

கூட்டுச் சராசரியைக் கணக்கிடும்போது அங்குள்ள எண்கள் மிகப்பெரிய எண்களாக இருந்தால் அவற்றிலிருந்து பொதுவான ஒரு எண்ணை (Assumed Mean அல்லது தோராய மதிப்பு) எல்லா எண்களிலிருந்தும் கழித்துவிட்டு சராசரி கண்டுபிடித்துப் பிறகு வந்த சராசரியுடன் கழித்த எண்ணைக் கூட்டி உண்மையான கூட்டுச் சராசரியைக் காணலாம். உதாரணத்திற்கு 1001, 1002, 1003 ஆகிய எண்களின் கூட்டுச் சராசரியைக் கணக்கிட 1, 2, 3 ஆகிய எண்களின் கூட்டுச் சராசரியைக் கண்டுபிடித்து (2) 1000உடன் கூட்டினாலே போதும் (1002). அதுபோல பொருள்களின் எடை கிராமில் கொடுக்கப்பட்டு எண்கள் பெரியதாக இருந்தால் (2010, 3100, 4005 என்பது போல) அவற்றைக் கிலோவாக மாற்றிச் சிறிய எண்களாக ஆக்கிவிடலாம். பிறகு கூட்டுச் சராசரி கணக்கிடுவது எளிதாகிவிடும்.

சில சமயம் கிடைக்கும் எண்களின் தொகுதியில் உள்ள ஆரம்ப எண்களும் முடிவு எண்களும் தெரியாமல் இருக்கலாம். உதாரணத்திற்கு அட்டவணை 11ஐப் பார்க்கவும்.

அட்டவணை - 11
மதிப்பெண்கள் விபரம்

மதிப்பெண்கள்	மாணவர்களின் எண்ணிக்கை
< 59	3
60 - 75	54
80 ≤	29

இதில் உள்ள புள்ளி விவரங்களுக்குக் கூட்டுச் சராசரி கண்டுபிடிக்க முடியாது. ஏனெனில், முதல் மற்றும் மூன்றாவது பிரிவுகளில் முறையே கீழ் எல்லையும் மேல் எல்லையும் இல்லை. எனவே அப்பிரிவுகளுக்கு மைய மதிப்பு (Mid value)

கண்டுபிடிக்க முடியாது. அதனால் fx கண்டுபிடிக்க முடியாது; அதனால் கூட்டுச் சராசரி கண்டுபிடிக்க முடியாது.

பல எண்களைக் கொண்ட ஒரு நீண்ட தொடருக்கு கூட்டுச் சராசரி கணக்கிடும்போது கணினியில் எண்களை நீண்ட நேரமாகப் பதிவு செய்ய வேண்டியிருக்கும். அல்லது ஒருவர் எண்களைச் சொல்ல இன்னொருவர் எண்களைக் கணினியில் பதிவு செய்ய வேண்டியிருக்கும். இப்படிபட்ட சந்தர்ப்பங்களில் கூட்டுச் சராசரி கண்டுபிடித்த பின்னர் தவறுகள் தெரிய வரலாம். உதாரணத்திற்கு 18 என்பதற்குப் பதிலாக 80 என்று பதிவு செய்து இருக்கலாம். அல்லது 30க்குப் பதிலாக 13 என்று பதிவு செய்து இருக்கலாம். இது பின்னர் கண்டுபிடிக்கப்பட்டால் கூட்டுச் சராசரி கண்டுபிடிப்பதற்கான எல்லாச் செயல்களையும் மீண்டும் தொடர வேண்டிய அவசியமில்லை. அந்த எண்களின் மொத்தத்திலிருந்து (Total) தவறான எண்களைக் கழித்துவிட்டு சரியான எண்களைக் கூட்டி புது சரியான மொத்தத்தைக் கண்டுபிடித்து அதிலிருந்து சரியான கூட்டுச் சராசரியைக் கண்டுபிடித்து விடலாம்.

அதுபோல தவறான அலைவெண்ணை மைய மதிப்புடன் பெருக்கியதால் தவறான கூட்டுச் சராசரி வரலாம். அப்படியிருந்திருந்தால், தவறான பெருக்கல் தொகையை (Mid value x f) மொத்தத்திலிருந்து கழித்துவிட்டு பின்னர் சரியான பெருக்கல் தொகையை மொத்தத்துடன் கூட்டி சரியான கூட்டுச் சராசரியைக் கணக்கிட்டு விடலாம்.

ஆய்வு செய்யும்போது ஓர் எண் தொகுதியினைப் பல குழுக்களாகப் பிரித்து பகுப்பாய்வு செய்தால் நிறையச் செய்திகள் கிடைக்கும். அப்படிச் செய்யும்போது ஒவ்வொரு குழுவுக்கும் ஒரு கூட்டுச் சராசரி கணக்கிட வேண்டி வரலாம். அதேசமயம் அந்த எண் தொகுதி முழுமைக்கும் ஒரு கூட்டுச் சராசரி கணக்கிடவும் தேவைப்படலாம். உதாரணத்திற்கு 1000

மாணவர்களை வகுப்புவாரியாக (5ஆம் வகுப்பு, 6ஆம் வகுப்பு, 7ஆம் வகுப்பு ...) பிரித்து அவர்களை ஒப்பிட்டுப் பார்க்க ஒவ்வொரு வகுப்பு மாணவர்களுக்கும் கூட்டுச் சராசரி கண்டுபிடிக்கலாம். பிறகு மொத்த மாணவர்களுக்கும் கூட்டுச் சராசரி வேண்டுமெனில் மீண்டும் 1000 மாணவர்களின் மதிப்பெண்களைக் கூட்டுவது கடினமான பணியாகிவிடும். அதற்குப் பதிலாக ஒவ்வொரு வகுப்புக் கூட்டுச் சராசரியையும் அந்தந்த வகுப்பில் உள்ள மாணவர்களின் எண்ணிக்கையால் பெருக்கி பின்பு அவற்றைக் கூட்டி 1000ஆல் வகுத்து விட்டால் மொத்த மாணவர்களுக்குமான கூட்டுச் சராசரி கிடைத்துவிடும்.

அலைவெண்கள் பொதுவாக முழு எண்களாகவே இருக்கும். ஆனாலும் அவை பின்னமாகவோ ஒன்றுக்கும் குறைந்து (less than 1) 0.4, 0.5 போன்றோ இருக்கவும் செய்யலாம். உதாரணத்திற்கு, நிகழ்தகவே அலைவெண்களாக இருக்கலாம். உதாரணத்திற்கு, இரண்டு நாணயங்களைச் சுண்டி விட்டால் அவை கீழே விழும்போது மேலே தெரிவது இரண்டு நாணயங்களிலும் தலையாகவோ, பூவாகவோ அல்லது ஒரு தலை ஒரு பூவாகவோ இருக்கலாம். இந்நிகழ்வுகளைக் கீழே வருமாறும் தரலாம். (நிகழ்தகவு எப்படி வந்தது என்பதை பின்னர் விளக்கமாகக் காணலாம்).

அட்டவணை - 12

இரண்டு நாணயங்கள் சுண்டப்படும்போது நேரிடக்கூடிய நிகழ்வுகளும் அவற்றிற்கான வாய்ப்புகளும்

தலையின் எண்ணிக்கை	அதற்கான நிகழ்தகவு	fx
0	¼	0
1	½	½
2	¼	½

அட்டவணை 12இல் உள்ள விபரங்களுக்கு fx கண்டுபிடித்து அவற்றைக் கூட்டினாலே கூட்டுச் சராசரி கிடைத்துவிடும். Σfஆல் Σxஐ வகுக்க வேண்டிய அவசியம் இல்லை. ஏனெனில், மொத்த நிகழ்தகவினையும் கூட்டினால் 1தான் வரும். எனவே Σfxஐ 1ஆல் வகுத்தாலும் வகுக்கா விட்டாலும் ஒரே விடைதான் வரும். கொடுக்கப்பட்டுள்ள உதாரணத்தில் (அட்டவணை 12ல்) $\Sigma fx = \frac{1}{2} + \frac{1}{2} = 1$. இதுதான் அட்டவணை 12இல் கொடுக்கப்பட்டுள்ள புள்ளி விபரங்களுக்கான கூட்டுச் சராசரியாகும்.

எடையிட்ட / நிறையிட்ட சராசரி (Weighted Mean)

சில சமயங்களில் மேலே விவரிக்கப்பட்ட கூட்டுச் சராசரி பொருந்தாது; எடையிட்ட கூட்டுச் சராசரி பொருந்தலாம். உதாரணத்திற்கு ஒரு மாணவர் ஒரு முதுகலைப் பட்டப்படிப்பில் நான்கு பருவங்களில் பெற்ற பருவ சராசரி மதிப்பெண்கள் 40, 45, 50, 65 என்று கொள்வோம். அப்படியானால் அந்த மாணவர் அந்த முதுகலைப் பட்டப் படிப்பில் பெற்ற சராசரி மதிப்பெண் என்ன என்று கேட்டால் *50* என்று (40 + 45 + 50 + 65 = 200; 200 ÷ 4 = 50) சொல்லலாம். இது ஒவ்வொரு பருவத்திலும் சமமான எண்ணிக்கையில் பாடங்கள் இருந்திருந்தால் மட்டுமே சரியாகும். அப்படியில்லாமல், முதலாம், இரண்டாம், மூன்றாம் மற்றும் நான்காம் பருவங்களில் வெவ்வேறு எண்ணிக்கையில் முறையே 2, 4, 6, 8 பாடங்கள் இருந்திருந்தால் மொத்தப் படிப்பிற்கான பாடச் சராசரி மதிப்பெண் {[(40 × 2) + (45 × 4) + (50 × 6) + (65 × 8)] ÷ 20} ஆகும். *(50 அல்ல).*

எடையிட்ட கூட்டுச் சராசரிக்கான சூத்திரம் = $\bar{X} = \frac{\Sigma WX}{\Sigma W}$

பெருக்குச் சராசரி

சராசரி கண்டுபிடிக்க வேண்டிய எண்கள் விகிதங்களாக (ratios) இருந்தால் பெருக்குச் சராசரி பொருத்தமாகும். உதாரணத்திற்கு, ஒரு கோப்பை காப்பியின் விலை ஒரு கோப்பை தேநீரின் விலையைப் போல் மூன்று மடங்காக 2008ஆம் ஆண்டு இருந்தது. இந்த விகிதம் 2009ஆம் ஆண்டு இரண்டு மடங்காக ஆகிவிட்டது. இதற்குச் சராசரி காண பெருக்குச் சராசரி முறைதான் பொருத்தம்.

காப்பியின் விலை *(3)* தேநீரின் விலை *(1)* → 2008இல்

காப்பியின் விலை *(2)* தேநீரின் விலை *(1)* → 2009இல்

இதற்குப் பெருக்குச் சராசரி இருவகையாகக் கணக்கிடலாம்.

1. தேநீர் விலையில் காப்பி விலை விகிதத்தின் சராசரி =

$$\sqrt{3 \times 2} = \sqrt{6} = 2.45 = \frac{1}{0.408}$$

2. காப்பி விலையில் தேநீர் விலை விகிதத்தின் சராசரி =

$$\sqrt{\frac{1}{3} \times \frac{1}{2}} = \sqrt{0.167} = 0.408 = \frac{1}{2.45}$$

இந்த மாதிரி தலைகீழிகள் சமமாக கூட்டுச் சராசரியில் வராது. செய்து பார்த்தால் $\frac{3+2}{2} = \frac{5}{2} = 2.5$ என்று முதலாவதும், $\left[\frac{1}{3} + \frac{1}{2}\right] \div 2 = \frac{5}{6} \div 2 = \frac{5}{12}$ என்று இரண்டாவதும் வரும். ஆனால் பெருக்குச் சராசரிக்கு வந்தது போல ஒன்றின் தலைகீழி இன்னொன்றுக்குச் சமமாக வராது. $\frac{1}{2.5} = 0.40$; $\frac{5}{12} = 0.42$ என்றுதான் வரும். அதுபோல $2.5 \neq \frac{12}{5}$. இவ்வாறு வித்தியாசங்கள் வருவதால், இந்த மாதிரி விகிதங்களைக் கொண்ட கணக்குகளுக்கு பெருக்குச் சராசரி மிகப் பொருத்தமாக அமைகிறது.

மேலும், வளர்ச்சி விகிதங்களின் (growth rates) சராசரி கண்டுபிடிக்கவும் கூட்டுச் சராசரி பொருந்தாது; பெருக்குச் சராசரி பொருத்தமாக அமையும். உதாரணத்திற்கு ஒரு கடல் மீனின் எடை மூன்று நாட்களில் 1000 கிலோவிலிருந்து 4000 கிலோவாகக் கூடியது என்றால், ஒரு நாளுக்கான சராசரி சதவீதக் கூடுதல் எவ்வளவு என்று கணக்கிடுவதற்கு 300% ÷ 3 = 100% என்று கொள்ளக்கூாடது. இது தவறான பதிலைத் தரும். ஏனெனில் 1000க்கு 100 சதவீத வளர்ச்சி முதல் நாள் முடிவில் 2000 கிலோவாகவும், பின்னர் இரண்டாம் நாள் முடிவில் 100 சதவீத வளர்ச்சி 4000 கிலோவாகவும், மூன்றாம் நாள் முடிவில் 100 சதவீத வளர்ச்சி 8000 கிலோவாகவும் வரும். ஆனால் உண்மையில் மூன்றாவது நாள் முடிவில் 4000 கிலோதான். இதில் உண்மையான சதவீத சராசரி 58.7% தான். இந்த 58.7 சதவீத சராசரி வளர்ச்சி வீதத்தை வைத்து கொடுத்த கணக்கை மீண்டும் செய்து பார்த்தால்,

முதல் நாள் முடிவில் 1000 + 587 = 1587

இரண்டாம் நாள் முடிவில் 1587 + 931.56 (1587ன் 58.7%) = 2518.56

மூன்றாம் நாள் முடிவில் 2518.56 + 1478.44 (2518.56ன் 58.7%) = 4000

இவ்வாறாகச் சராசரி சதவீத வளர்ச்சி 58.7% என்பதுதான் சரி. அவ்வாறின்றி, 4000 - 1000 = 3000. இது 300 சதவீத வளர்ச்சி; இதனுடைய சராசரி $\frac{300}{3}$ = 100% என்று கொள்ளக்கூடாது. இந்தக் கணக்கில் வந்துள்ள 58.7% எப்படிக் கண்டுபிடிக்கப் பட்டது என்பதற்குக் கீழ்வரும் படிகளைக் கவனிக்க வேண்டும்.

முதல் நாள் முடிவில் : $1000 + 1000r = 1000\ (1 + r)$

இதில் r என்பது வளர்ச்சி வீதம் (growth rate)

இரண்டாம் நாள் முடிவில் :

$$1000(1 + r) + 1000\ (1 + r)r = 1000\ (1 + r)^2$$

மூன்றாம் நாள் முடிவில் :

$$1000(1 + r)^2 + 1000\,(1 + r)^2 r = 1000\,(1 + r)^3$$

இதில் $1000\,(1 + r)^3 = 4000$ எனில்,

$$(1 + r)^3 = 4;\ (1 + r) = \sqrt[3]{4};\ r = \sqrt[3]{4} - 1$$

$= 1.587 - 1 = 0.587$ சதவீதத்தில் 58.7%

பல எண்கள் உள்ளடக்கிய கணக்காக இருந்தால் கணினிகளும் கணக்கிகளும் இல்லாத நாட்களில் அதற்கு வர்க்கமூலம் காண்பது சிரமமாக இருந்தது. எனவே அந்த எண்களை லாகிருத (Logarithm) எண்களாக மாற்றி அவற்றைக் கூட்டி பின்னர் அந்த மொத்தத்தை அத்தொடரில் உள்ள எண்களின் எண்ணிக்கையால் வகுத்துப் பிறகு அதற்கு எதிர் லாகிருத (antilogarithm) எண்ணைக் கண்டுபிடித்தால் அது பெருக்குச் சராசரியாக இருக்கும். ஆனால், தற்போது பல மடங்கு வர்க்கமூலங்கள் கண்டுபிடிப்பது கணக்கி மற்றும் கணினி உதவி மூலம் எளிதாகிவிட்டதால், லாகிருத அடிப்படையிலான முறை தேவையில்லாமல் போய்விட்டது. இருப்பினும் அதையும் தெரிந்து கொள்வது நல்லதே. எனவே, கீழ்வரும் எடுத்துக்காட்டு தரப்படுகிறது.

331, 411, 251, 713, 812 ஆகிய எண்களின் பெருக்குச் சராசரி காண, முதலில் அவற்றை பின்வருவதுபோல் லாகிருத மதிப்புகளுக்கு மாற்றி அவற்றைக் கூட்டி ஐந்தால் (இங்கு ஐந்து எண்கள் தரப்பட்டுள்ளதால்) வகுத்து விடவேண்டும்.

$$2.51 + 2.62 + 2.40 + 2.85 + 2.91 = 13.29$$

$$\frac{13.29}{5} = 2.66$$

இதனுடைய (2.66) எதிர் லாகிருத மதிப்பான 456.2 என்பதுதான் இத்தொடரின் பெருக்குச் சராசரியாகும்.

இங்கு அலைவெண்கள் (frequency) கொடுக்கப்பட்டு இருந்தால், லாகிருத மதிப்புக்கள் கண்டுபிடித்த பின்னர், அவற்றிற்குரிய அலைவெண்களால் பெருக்கி, அவற்றின் மொத்தம் கண்டுபிடித்து அந்த மொத்தத்தை அலைவெண்களின் மொத்தத்தால் வகுத்து வந்த ஈவுக்கு எதிர்லாகிருதம் கண்டுபிடிக்க வேண்டும். கீழேகொடுக்கப்பட்டுள்ள கணக்கைப் பார்க்கலாம்.

அட்டவணை - 13

லாகிருத முறைப்படி பெருக்குச் சராசரி

x	f	log x	f log x
2	4	0.30	1.20
5	9	0.69	6.29
6	11	0.78	8.56
8	6	0.91	5.42
மொத்தம்	**30**		**21.47**

$$\frac{\Sigma f \log x}{\Sigma f} = \frac{21.47}{30} = 0.72$$

0.72 ன் எதிர்லாகிருத மதிப்பு 5.19.

எனவே, பெருக்குச் சராசரி 5.19 ஆகும்.

கொடுக்கப்பட்டுள்ள அட்டவணை 14இல் இருந்து ஒவ்வொரு 10 ஆண்டுகளிலும் சராசரி சதவீத மக்கள்தொகை அதிகரிப்பை எவ்வாறு கணக்கிடலாம் என்று பார்க்கலாம்.

அட்டவணை - 14
ஆண்டுவாரியாக மக்கள்தொகை

ஆண்டு	மக்கள்தொகை (கோடியில்)
1971	250.2
1981	279.6
1991	283.9
2001	303.4

மேலே அட்டவணையில் கொடுக்கப்பட்டுள்ள மக்கள்தொகை எண்களின் சார்பு மதிப்புக்களைக் கீழ்க்கண்டவாறு கணிக்கலாம்.

முதலில் ஒவ்வொரு ஆண்டின் மக்கள்தொகையும் (கொடுக்கப்பட்டுள்ள ஆண்டுகளில்) அதற்கு முந்தைய ஆண்டின் மக்கள்தொகையை ஒப்பிடுகையில் எவ்வளவு வளர்ந்துள்ளது எனக் கணக்கிடலாம். முதலில் 1981க்குக் கணக்கிடலாம்.

250.2 கோடி 279.6 கோடியானால்

100 எவ்வளவாகும்?

$$\frac{279.6 \times 100}{250.2} = 111.8 \text{ ஆகும்.}$$

இதேபோல் மற்ற ஆண்டுகளுக்கும் கணக்கிட்டால், அவை 1991க்கு 101.6 ஆகவும், 2001க்கு 106.9 ஆகவும் இருக்கும். அவற்றின் லாகிருத எண்களின் கூடுதல் 2.05 + 2.00 + 2.03 = 6.08. இதை மூன்றால் வகுத்தால் 2.028. இதற்கு எதிர் லாகிருத மதிப்பு 106.7. இதிலிருந்து 100ஐக் கழித்து, 6.7 தான் சராசரி சதவீத மக்கள்தொகை அதிகரிப்பு.

இந்தச் சராசரி சதவீத மக்கள்தொகை அதிகரிப்பைப் பயன்படுத்தி ஒவ்வொரு 10 ஆண்டுகளிலும் மக்கள்தொகை வளர்ந்த விதத்தைக் கணக்கிடலாம்.

1971இல் 250.2

1981இல் 250.2 + 250.2 (0.067) = 266.9

1991இல் 266.9 + 266.9 (0.067) = 284.8

2001இல் 284.8 + 284.8 (0.067) = 303.4

இதற்குப் பதிலாக (303.4 - 250.2) ÷ 3 = 17.73 என்று சராசரி கணக்கிடுவது தவறாகும்.

இசைவுச் சராசரி (Harmonic Mean)

சராசரி வேகம், சராசரி நேரம் போன்றவற்றைக் கணக்கிடுவதற்கு இசைவுச் சராசரி பொருத்தமானதாகும். ஒரு வானவூர்தி ஒவ்வொரு பக்கத்தின் நீளமும் *100* கிலோ மீட்டர் *(கி.மீ.)* தூரமுள்ள நான்கு பக்கங்களைக் கொண்ட சதுர வடிவமான பாதையில் பறக்கும்பொழுது முதல் பக்கத்தை மணிக்கு *100 கி.மீ.* வேகத்திலும், இரண்டாவது, மூன்றாவது மற்றும் நான்காவது பக்கங்களை முறையே மணிக்கு *200கி.மீ., 300 கி.மீ., 400 கி.மீ.* வேகத்திலும் கடந்து சென்றால், அதன் சராசரி வேகத்தைக் கணக்கிட இசைவுச் சராசரி முறையைப் பயன்படுத்தலாம்.

இசைவுச் சராசரியை H என்றும், வேகத்தை X என்றும் பாதைகளின் எண்ணிக்கை *(4)*யை n என்றும் Σ கூடுதல் குறியென்றும் கொண்டு,

$$H = \frac{n}{\Sigma \frac{1}{X}}$$ என்று கொள்ளலாம்.

அலைவெண்கள் (frequency) இருந்தால்

$$H = \frac{\Sigma f}{\Sigma \frac{f}{X}}$$ என்று கொள்ளலாம்.

தொடர்ச்சியான அலைவெண் பரவலுக்கு ஒவ்வொரு பிரிவு (Class)க்கும் பிரிவு மையம் (Mid Value) கண்டுபிடித்து அவற்றை X ஆகக் கொண்டு Hஐக் கணிக்கலாம்.

சில சராசரி முறைகள்

முறை 1

ஒருவர் ஒரு பொருளை கிலோ ஒன்றுக்கு ரூ.16, ரூ.18, ரூ.21, ரூ.25 என நான்கு நாட்கள் வாங்கியிருந்தால், அந்தப் பொருளின் சராசரி விலை கிலோ ஒன்றுக்கு எவ்வளவு? இதற்கு இசைவுச் சராசரி முறையைப் பயன்படுத்தலாம்.

அந்தப் பொருளின் சராசரி விலை கிலோ ஒன்றுக்கு

$$= \frac{4}{\frac{1}{16}+\frac{1}{18}+\frac{1}{21}+\frac{1}{25}}$$

$$= \frac{4}{\frac{9450+8400+7200+6048}{151200}} = \frac{4}{\frac{31098}{151200}}$$

$$= 4 \div \frac{31098}{151200} = 4 \times \frac{151200}{31098} = \text{ரூ. } 19.45$$

முறை 2

ஒருவேளை அவர் அந்தப் பொருளுக்காகத் தினமும் ரூ.200 செலவு செய்திருந்தால், சராசரி செலவு கிலோ ஒன்றுக்கு

$$= \frac{200 \times 4}{12.5 + 11.11 + 9.52 + 8} = \frac{\text{மொத்த செலவு}}{\text{மொத்தப் பொருள்கள்}}$$

$$= \frac{800}{41.13} = \text{ரூ. } 19.45$$

மேற்காணும் இரு முறைகளிலும் சராசரி விலை / செலவு ஒன்றுக்குச் சமமாக வந்துள்ளது. இதில் முதல் முறை இசைவுச் சராசரியாகவும் இரண்டாவது முறை கூட்டுச் சராசரியாகவும் உள்ளதைக் கவனிக்கலாம்.

முறை 3

ஒருவேளை அவர் அந்தப் பொருளை ஒவ்வொரு நாளும் 1000 எண்ணிக்கையில் வாங்கியிருந்தால் (அதே விலைகளில்), அப்பொழுது சராசரிச் செலவு எப்படிக் காண்பது?

$$\text{கிலோ ஒன்றுக்கு சராசரி செலவு} = \frac{\text{மொத்த செலவு}}{\text{பொருள்களின் மொத்த எண்ணிக்கை}}$$

$$= \frac{16000 + 18000 + 21000 + 25000}{4000} = \frac{80000}{4000} = \text{ரூ. } 20$$

இந்த மூன்று முறைகளுக்கும் இடையேயுள்ள ஒற்றுமை வேற்றுமைகளை யோசித்துப் பார்த்தால் மேலும் சில சுவையான உண்மைகள் தெரியவரும்.

இருபடிச் சராசரி அல்லது வர்க்கமூலச் சராசரி (Quadratic Mean or Root Mean Square - RMS)

கொடுக்கப்பட்டுள்ள எண்களுக்கு வர்க்கம் கண்டுபிடித்து அதைக்கூட்டி அந்த எண்களின் எண்ணிக்கையால் வகுத்து பின்னர் அதற்கு மூலம் கண்டுபிடித்தால் கிடைப்பதற்குப் பெயர் வர்க்கமூலச் சராசரி (அதாவது square root of the mean of the squared values)

உதாரணத்திற்கு *1, 2, 4, 6* என்ற எண்களின்

$$\text{RMS} = \sqrt{\frac{1^2 + 2^2 + 4^2 + 6^2}{4}} = \sqrt{\frac{1 + 4 + 16 + 36}{4}} = \sqrt{14.25} = 3.77$$

இந்த எண்களுக்கு கூட்டுச் சராசரி 3.25ஆக $(\frac{13}{4})$ இருந்திருக்கும்.

சராசரிகளுக்கிடையேயான உறவு

கூட்டுச் சராசரியை AM என்றும் பெருக்குச் சராசரியை GM என்றும் இசைவுச் சராசரியை HM என்றும் கொள்வோமே யானால், AM ≥ GM ≥ HM. ஒரு தொடரில் உள்ள எண்கள் அனைத்தும் சமமான (அல்லது ஒரே) எண்ணாக இருந்தால்

AM = GM = HM. உதாரணத்திற்கு 2, 4, 8 ஆகிய எண்களின் AM = 4.67, GM = 4, HM = 3.43. அதேசமயம் 2, 2, 2 ஆகிய எண்களின் AM = 2, GM = 2, HM = 2. ஒரு தொடரில் உள்ள எண்களில் ஏதாவது ஒன்றோ அல்லது ஒன்றுக்கு மேற்பட்டோ பூஜ்யம் (0) இருந்தால் GM பூஜ்யமாகிவிடும். ஆனால் மற்ற சராசரிகள் அப்படி ஆவதில்லை. கொடுக்கப்பட்ட எண்களில் ஓர் எண் எதிர்மறை எண்ணாக (negative) இருந்தால் GM காணமுடியாது. எனவே அந்த மாதிரி சூழ்நிலைகளில் GM உதவாது.

இடைநிலை

கொடுக்கப்பட்டுள்ளவைகளை ஏறு வரிசையிலோ இறங்கு வரிசையிலோ வரிசைப்படுத்திய பின்பு, அவ்வரிசையில் மையமாக அமைந்துள்ளது இடைநிலை எனப்படுகிறது. மூன்று வேலையாட்களின் கூலி ரூ.60, ரூ.70, ரூ.65 ஆக இருந்தால், அவற்றை ஏறு வரிசையிலோ இறங்கு வரிசையிலோ வரிசைப்படுத்தி அவற்றிற்கு நடுவில் (65) இருப்பதை இடைநிலையாகக் கொள்ளலாம். இங்கு மூன்று வேலையாட்கள் (ஒற்றைப்படை எண்) ஆதலால் (3+1) ÷ 2 என்று கொண்டு 60, 65, 70இல் இரண்டாவது உள்ளதுதான் இடைநிலை எனக் குறிக்கப்படுகிறது. நான்கு (இரட்டைப் படை எண்) கூலியாட்களாக இருந்தால் (4 + 1) ÷ 2 = 2.5 என்று கொண்டு இரண்டாவது எண்ணுக்கும் மூன்றாவது எண்ணுக்கும் இடையில் உள்ள எண் இடைநிலையாகும். உதாரணத்திற்கு, நான்கு வேலையாட்களின் கூலி ரூ.60, ரூ.70, ரூ.66, ரூ.72 என இருந்தால் 60, 66, 70, 72 என ஏறு வரிசையில் எழுதி (அல்லது 72, 70, 66, 60 என இறங்கு வரிசையில் எழுதி) [66 + 70] ÷ 2 = 68ஐ இடைநிலையாகக் கொள்ளலாம்.

தொடர்ச்சியற்ற அலைவெண்பரவல் கொடுக்கப் பட்டிருந்தால் கீழ்வருமாறு இடைநிலை காணலாம்.

அட்டவணை - 15

கூலி மற்றும் கூலியாட்களின் எண்ணிக்கை

கூலி	கூலியாட்களின் எண்ணிக்கை
50	2
60	5
65	8
70	4
மொத்தம்	19

இந்த அலைவெண்களைப் பயன்படுத்தி குவிவு அலைவெண்களைக் கணக்கிட வேண்டும். அவை 2, 7, 15, 19 என்று அமையும். இது ஒற்றைப்படை எண். (19+1) ÷ 2 = 10. எனவே 10ஆவது இடத்தில் வரக்கூடியது இடைநிலையாகும். அதன்படி பார்த்தால், 65 இடைநிலையாகும். இதைத்தான் பத்தாவது கூலியாள் கூலியாகப் பெற்றிருப்பார்.

அட்டவணை - 16

மாணவர்களின் உயரம்

உயரம் (செ.மீ.)	மாணவர்களின் எண்ணிக்கை
117.5 - 126.5	3
126.5 - 135.5	5
135.5 - 144.5	9
144.5 - 153.5	12
153.5 - 162.5	5
162.5 - 171.5	4
171.5 - 180.5	2

இதற்கு குவிவு அலைவெண்கள் 3, 8, 17, 29, 34, 38, 40 ஆகும். இரண்டு முறைகள் மூலம் இடைநிலை காணலாம்.

முறை 1 : இடைச்செருகல் முறை

இங்கு 40 மாணவர்கள் இருப்பதால் 20ஆவது மாணவரின் உயரத்திற்கும் 21ஆவது மாணவரின் உயரத்திற்கும் இடைப்பட்டது இடைநிலையாக இருக்கும். முதல் 17 மாணவர்களின் உயரம் 144.5 சென்டி மீட்டருக்குள் உள்ளது. எனவே 20ஆவது மாணவரின் உயரம் 144.5 செ.மீட்டருக்கு மேலும் 153.5 செ.மீட்டருக்குள்ளும் இருக்க வேண்டும். 144.5செ.மீ.க்கும் 153.5 செ.மீ.க்கும் இடையில் 12 மாணவர்கள் உள்ளனர். இந்தப் பன்னிரண்டு மாணவர்களில் முதல் மூன்று மாணவர்களின் உயரம் $144.5 + \frac{3}{12}(153.5 - 144.5)$ ஆக இருக்கும். அதாவது 146.8 செ.மீ. ஆக இருக்கும்.

முறை 2 : சூத்திரத்தைப் பயன்படுத்தும் முறை

20ஆவது மாணவர் நான்காவது பிரிவில் இருப்பார். ஏனெனில், முதல் மூன்று பிரிவிற்குள் 17 மாணவர்களே உள்ளனர். எனவே, நான்காவது பிரிவின் கீழ் எல்லை (L_1) அதாவது 144.5. அந்தப் பிரிவின் அலைவெண் (f), அதாவது 12. அந்தப் பிரிவிற்கு முந்தைய பிரிவின் குவிவு அலைவெண் (m), அதாவது 17. n/2 = 40/2. பிரிவு இடைவெளி (c) 9. இந்த விபரங்களைப் பயன்படுத்தி இடைநிலையைக் காணலாம்.

$$\text{இடைநிலை} = l_1 + \frac{n/2 - m}{f} \times c$$

$$= 144.5 + \left(\frac{20-17}{12}\right) \times 9 = 146.8 \text{ செ.மீ.}$$

பரவல் செவ்வக வரைபடம் மூலமும் மேலின கீழின குவிவு அலைவெண்கள் மூலம் (முன்னரே விளக்கியது போல) இடைநிலையைக் காணலாம்.

மேல் எல்லையும் கீழ் எல்லையும் கொடுக்கப்படாத தொடருக்கு இடைநிலை பொருத்தமாகும்; கூட்டுச்சராசரி காண இயலாது.

இடைநிலை கொடுக்கப்பட்ட எல்லா எண்களுக்கும் தொடர்பில்லாததாக இருக்கும். ஒரு தொடரில் முதல் அல்லது இறுதியில் உள்ள சில எண்களை மாற்றினாலோ நீக்கினாலோ கூட இடைநிலை மாறாமல் இருக்கலாம். ஆனால் கூட்டுச் சராசரி அப்படியல்ல. சிறுமாற்றம் செய்தால் கூட கூட்டுச் சராசரி மாறிவிடும் தன்மையுடையது.

கால்மானங்கள் (Quartiles)

இடைநிலை ஒரு தொடரை இரண்டாகப் பிரிக்கும் மையப்பகுதியாக விளங்குகிறது. கொடுக்கப்பட்டுள்ள தொடரை நான்கு சம பகுதிகளாகப் பிரிப்பது கால்மானம் ஆகும். எனவே முதல் கால்மானம் இடைநிலையில் பாதியாகவும், இரண்டாம் கால்மானம் இடைநிலைக்குச் சமமாகவும் இருக்கும். மூன்றாவது பகுதியையும் நான்காவது பகுதியையும் பிரிப்பது மூன்றாவது கால்மானம் (upper quartile) ஆகும்.

முதல் கால்மானம் (lower quartile or first quartile) Q_1 என்றும் இரண்டாம் கால்மானம் (second quartile) Q_2 என்றும் மூன்றாம் கால்மானம் (Third quartile) Q_3 என்றும் குறிக்கப்படுகின்றன. முதல் கால்மானம் கண்டுபிடிக்க $\frac{N}{4}$ என்றும், இரண்டாம் கால்மானத்திற்கு $\frac{N}{2}$ என்றும், மூன்றாம் கால்மானத்திற்கு $\frac{3N}{4}$ என்றும் கணக்கிட வேண்டும். மற்ற செய்முறைகள் சிறிய பொருத்தமான மாற்றங்களுடன் இடைநிலைக்கு உள்ளவைபோல்தான்.

இதேபோல் ஒரு பரவலை ஐந்தாகப் (Quintiles) பிரிக்கலாம். அப்படியானால் ஒவ்வொரு பகுதியும் 20 சதவீதத்தைக் கொண்டிருக்கும். ஒரு பரவலை பத்துப்

பகுதிகளாகப் பிரித்தால் பதின்மானங்கள் (Deciles) என்றும் நூறு பகுதிகளாகப் பிரித்தால் நூற்றுமானங்கள் (Percentiles) என்றும் அழைக்கப்படுகின்றன. இடைநிலைக்கு உள்ள செய்முறைகளைச் சிறிது தக்கவாறு மாற்றி ஐம்மானங்கள் (quintiles), பதின்மானங்கள் (deciles) மற்றும் நூற்றுமானங்கள் (percentiles) கணக்கிடப்படுகின்றன.

முகடு (Mode)

கொடுக்கப்பட்டுள்ள எண்களின் தொடரில் எந்த மதிப்பு அதிகமான தடவைகள் வருகின்றதோ அந்த மதிப்பே முகடு எனப்படும். ஒரு தொடரில் ஒரு முகடோ (unimodal) இரு முகடுகளோ (Bimodal) இருக்கலாம். அதுபோல முகடு இல்லாமல் கூட இருக்கலாம்.

கீழ்வரும் தொடரில் முகடு இல்லை.

1, 3, 7, 16, 25, 34.

கீழ்வரும் தொடரில் ஒரே முகடு. அது 2 (Unimodal)

2, 2, 2, 3, 7, 8, 10, 15

கீழ்வரும் தொடரில் இரு முகடுகள் உள்ளன. (2,7)

2, 2, 2, 7, 7, 7, 8, 10, 15

ஒரு வரைபடத்தில் எந்த மதிப்புக்கு அலைவெண் வரைகோடு மிக உயரமாக உள்ளதோ, அம்மதிப்புதான் முகடு.

தொடர் மதிப்பு கொண்ட மாறிக்கு (Variable) கீழ்வரும் சூத்திரம் பயன்டுத்தப்படுகிறது.

$$\text{முகடு} = L_1 + \left(\frac{\Delta_1}{\Delta_1 + \Delta_2}\right)C$$

L_1 = முகட்டுப் பிரிவின் கீழ் எல்லை

Δ_1 = முகட்டுப் பிரிவின் அலைவெண்ணுக்கும் அதற்குக் கீழ் உள்ள பிரிவின் அலைவெண்ணுக்கும் உள்ள வித்தியாசம்

Δ_2 = முகட்டுப் பிரிவின் அலைவெண்ணுக்கும் அதற்குக் மேலே உள்ள பிரிவின் அலைவெண்ணுக்கும் உள்ள வித்தியாசம்

C = பிரிவு இடைவெளி (Class interval)

அட்டவணை 16இல் உள்ள விபரங்களைக் கொண்டு கீழேவருமாறு முகடு கணக்கிடலாம்.

$$\text{முகடு} = 144.5 + \left(\frac{3}{3+7}\right)9$$

$$= 147.2 \text{ செ.மீ.}$$

முகடு கண்டுபிடிக்க தொகுப்பு அட்டவணை முறையும் பயன்படுத்தப்படுகிறது. அது மிக நீண்ட செய்முறைகளைக் கொண்டதால், அந்த முறை இங்கு தவிர்க்கப்படுகிறது. அலைவெண்களைக் கொண்டு செவ்வக வரைபடம் மூலமும் முகட்டினைத் தீர்மானிக்கலாம்.

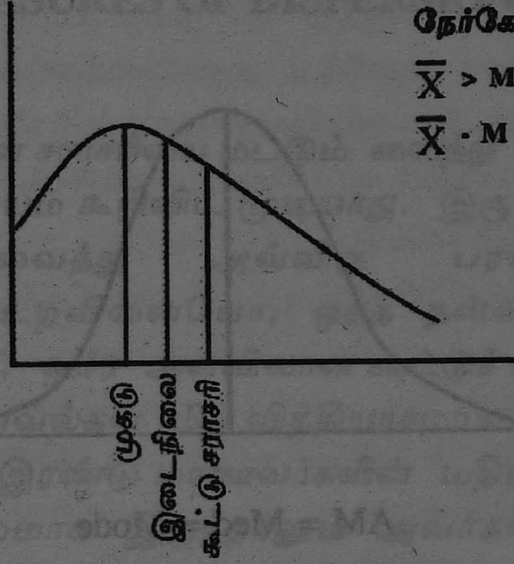
கூட்டுச் சராசரி, இடைநிலை மற்றும் முகடு ஆகியவற்றிற்கிடையேயான தொடர்பு

ஒரு முகட்டினைக் கொண்ட ஓரளவுக்கு சீரற்ற (skewed or asymmetrical) அலைவெண் வளைகோட்டுக்கு கீழ்க்காணும் உறவு கண்டுபிடிக்கப்பட்டுள்ளது.

(கூட்டுச் சராசரி - முகடு) = 3 (கூட்டுச் சராசரி - இடைநிலை)

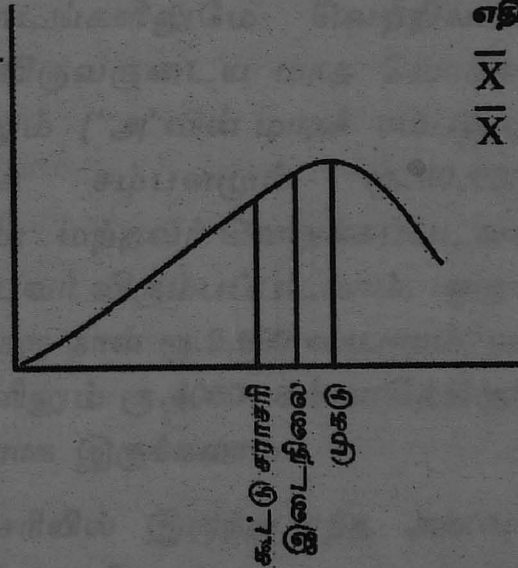
வரைபடம் - 12

வலதுபுறம் சிரற்ற பரவலுக்கான வரைபடம் (+)

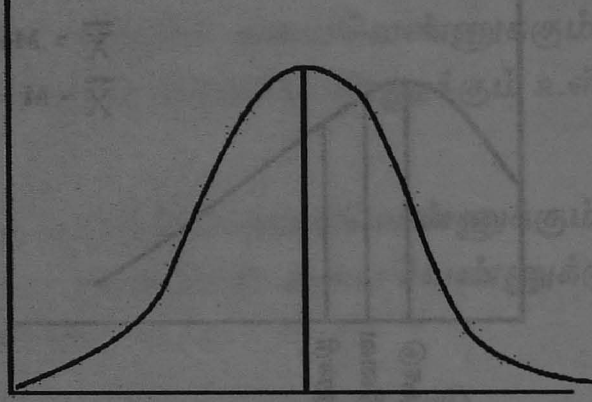


வரைபடம் - 13

இடதுபுறம் சிரற்ற பரவலுக்கான வரைபடம் (–)



வரைபடம் - 14



AM = Med = Mode

சீரான பரவலுக்கு (Symmetrical or normal) கூட்டுச்சராசரி (AM) இடைநிலை (Med) முகடு (M) ஆகியவை சமமாக இருக்கும்.

3. சிதறல் அளவைகள்
(MEASURES OF DISPERSION)

ஒரு பரவலின் சராசரியை மட்டும் வைத்து அப்பரவலின் முழுத்தன்மையையும் கூறிவிட முடியாது. இரு பரவல்களின் சராசரிகளை வைத்து அவ்விரு பரவல்களுக்கும் இடையேயுள்ள உறவினையோ, ஒத்த தன்மைகளையோ கூறவும் முடியாது. ஒரே அளவிலான கூட்டுச் சராசரிகளைக் கொண்ட இரு பரவல்கள் மிக வித்தியாசமாக இருக்கலாம். உதாரணத்திற்கு, இரண்டு மாணவர்களின் மதிப்பெண்களின் கூட்டுச்சராசரி சமமாக இருந்தாலும் அவர்களின் தேர்ச்சி நிலைமை வேறு மாதிரியாக இருக்கலாம். அன்பு மற்றும் ஆதி ஆகியவர்கள் இரண்டு பாடங்களில் பெற்ற மதிப்பெண்கள் முறையே 0, 100 மற்றும் 50, 50. அவர்கள் இருவரின் மதிப்பெண்களின் மொத்தம் 100 தான். கூட்டுச்சராசரியும் 50 தான். ஆனால், அன்பு ஒரு பாடத்தில் தோல்வி அடைந்துள்ளார். ஆதி இரண்டு பாடங்களிலுமே வெற்றியடைந்துள்ளார். அதுபோல், அவ்விருவருடைய மாத மொத்த வருமானம் சமமாக இருந்தாலும் ('அ' வின் மாதச் சம்பளமும் ரூ.60,000, 'ஆ' வின் மாதச் சம்பளமும் ரூ.60,000) ஒருவர் விரும்பப்படலாம்; மற்றவர் வெறுக்கப்படலாம். தினமும் ரூ.2,000 சம்பாதிப்பவர் விரும்பப்படலாம். ஒரு நாள் ரூ.5,000 நட்டமும், மற்றொரு நாள் ரூ.9,000 லாபமும் சம்பாதிப்பவர் (இரண்டு நாட்களிலும் ரூ.4000 சம்பாதித்திருந்த போதும்) விரும்பப்படாதவராக இருக்கலாம்.

கூட்டுச் சராசரியில் இருந்து எந்த அளவுக்கு எண்கள் விலகியுள்ளன என்பது சிதறல் என அழைக்கப்படுகிறது.

கூட்டுச்சராசரி, இடைநிலை, முகடு, பெருக்குச் சராசரி மற்றும் இசைவுச் சராசரி ஆகியவை முதல் நிலைச் சராசரிகள் (averages of the first order) என அழைக்கப்படுகின்றன. முதல் நிலைச் சராசரிகளிலிருந்து கொடுக்கப்பட்டுள்ள இனங்களின் விலக்கங்களைக் கணித்து அவ்விலக்கங்களின் சராசரியைச் சிதறல் அளவையாகக் கொள்வதால், சிதறல் அளவைகள் இரண்டாம் நிலைச் சராசரிகள் (averages of second order) என அழைக்கப்படுகின்றன.

ஒரு மாதிரியிலிருந்து கணிக்கப்பெற்ற சராசரி மதிப்பானது முழுமைத் தொகுதியின் சராசரி மதிப்பிற்கு சிறந்த மதிப்பீடாக அமையுமா இல்லையா என்பதைக் கண்டறிய சிதறல் அளவைகள் பயன்படுத்தப்படுகின்றன. மாறியின் மதிப்புகளுக்கிடையேயுள்ள வேறுபாட்டின் அளவையறிந்து அவ்வேறுபாட்டின் அளவைக் குறைப்பதற்கும் சிதறல் அளவைகள் பயன்படுகின்றன. ஒரு மாறியின் மதிப்புகளுக்கிடையே அதிக வேறுபாடுகள் இருந்தால் சிதறல் அளவை அதிகமாகவும், குறைவான வேறுபாடுகள் இருந்தால் சிதறல் அளவை குறைவாகவும் இருக்கும். உதாரணத்திற்கு, ஒரு வகுப்பில் உள்ள அனைத்து மாணவர்களும் சமமான மதிப்பெண்கள் பெற்றிருந்தால் சிதறல் அளவை பூஜ்யமாக (0) இருக்கும். வேறுபாடு கூடக்கூட சிதறல் அளவையும் பெரியதாகிக் கொண்டே போகும். இதற்கு அதிகப்படியான ஒரு வரம்பினைக் குறிப்பிட்டுக் கூறமுடியாது.

மிக அதிகமாகப் பயன்படுத்தும் சிதறல் அளவையாக திட்டவிலக்கத்தினைக் (standard deviation) கூறலாம். இது தவிர, வீச்சு (range) கால்மான விலக்கம் (quartile deviation), சராசரி விலக்கம் (mean deviation), லாரன்ஸ் வளைகோடு (Lorenz Curve) ஆகியவையும் உள்ளன. பொருளியல் மாறிகளான (Economic variables) வருமானம், செலவு, நிலம் ஆகியவை எவ்வாறு பகிர்ந்தளிக்கப்பட்டுள்ளன என்பதை லாரன்ஸ் வளைகோடு கொண்டு காட்டலாம்.

ஒவ்வொரு சிதறல் அளவையையும் முழுமையான அளவையாகவும் தரலாம்; ஒப்பீட்டு அளவை (relative measure - coefficient)யாகவும் தரலாம். ஒப்பீட்டு அளவைக்கு அலகுகள் (eg. kg, litre) பயன்படுத்தலாகாது.

வீச்சு

பல எண்கள் கொண்ட ஒரு தொடரில் மிகப்பெரிய எண்ணுக்கும் மிகச்சிறிய எண்ணுக்கும் உள்ள வித்தியாசம் (difference) வீச்சு என அழைக்கப்படுகிறது. 2, 3, 5, 8, 10, 12 என்ற எண்களின் வீச்சு 2 முதல் 12 வரை என்றோ, 2-12 என்றோ அழைக்கப்படலாம். கணக்கிட்டுக் கூற வேண்டுமானால் 12-2=10 என்று கூறலாம். இது ஒரு முழுமையான அளவையாகும் (absolute measure). மேலே கொடுக்கப்பட்டது கிலோவாக இருந்தால் இந்த அளவையையும் கிலோவில் குறிக்க வேண்டும். எனவே வெவ்வேறு அலகுகளில் கொடுக்கப்பட்டிருக்கும் இரண்டு வேறுபட்ட பரவல்களின் சிதறல்களை ஒப்பிடுவது சரியாகாது.

உதாரணத்திற்கு, ஒரு வகுப்பில் உள்ள மாணவர்களின் வயது ஆண்டுகளிலும் உயரம் சென்டி மீட்டரிலும் கொடுக்கப்பட்டிருந்தால் அவ்விரு மாறிகளையும் வீச்சை வைத்து ஒப்பிடலாகாது.

$$\text{வீச்சின் கெழு} = \frac{\text{பெரிய மதிப்பு} - \text{சிறிய மதிப்பு}}{\text{பெரிய மதிப்பு} + \text{சிறிய மதிப்பு}}$$

சராசரி விலக்கம் (Mean deviation or Average deviation)

$$\text{சராசரி விலக்கம்} = \text{M.D.} = \frac{\Sigma |X-\bar{X}|}{n}$$

இதில் X என்பது மாறியின் மதிப்பு, $\bar{X}$ என்பது கூட்டுச் சராசரி, n என்பது அத்தொடரில் உள்ள எண்களின் எண்ணிக்கை. | | என்ற இரண்டு செங்குத்துக் கோடுகள் அங்கு வரும் அனைத்து எண்களையும் (கழித்தல் அல்லது

எதிரிடையையும்) நேரிடையாகப்(கூட்டல்) பயன்படுத்த வேண்டும் எனக் குறிக்கின்றன. கூட்டுச் சராசரிக்கும் அந்தச் சராசரி கண்டுபிடிக்கப்பட்ட ஒவ்வொரு எண்ணுக்கும் உள்ள வித்தியாசத்தை அப்படியே (எதிரிடைக் குறியை அல்லது கழித்தல் அடையாளத்தை நேரிடைக்குறியாக அல்லது கூட்டல் அடையாளமாக மாற்றாமல்) கூட்டினால் அந்தக் கூடுதல் பூஜ்யத்தைத் (0) தரும் என்பதால் எதிரிடைக் குறியும் (கழித்தல் அடையாளமும்) நேரிடைக்குறியாக (கூட்டல் அடையாளமாக) மாற்றப்பட்டு வித்தியாசங்கள் கூட்டப்படுகின்றன.

2, 3, 6, 8, 11 ஆகிய எண்களின் சராசரி விலக்கத்தைக் காண முதலில் கூட்டுச் சராசரி காணவேண்டும். அது 6. பிறகு ஒவ்வொரு எண்ணிலிருந்தும் 6ஐக் கழித்து அதிலிருந்து கிடைக்கும் எண்களின் அடையாளங்களைக் கூட்டல் அடையாளங்களாகக் கொண்டு அவற்றினைக் கூட்ட வேண்டும்; அது 14.

எனவே, சராசரி விலக்கம் $= \frac{14}{5} = 2.8$

அங்கு அலைவெண்கள் (frequencies = f) இருந்தால்

$$MD = \frac{\Sigma f|X-\bar{X}|}{n} \; ; \quad n = \Sigma f.$$

சில சமயங்களில் கூட்டுச்சராசரிக்குப் பதிலாக இடைநிலையையோ அல்லது வேறு எந்த மையப்போக்கு அளவையோ கொண்டும் இந்தச் சராசரி விலக்கம் கண்டுபிடிக்கப்படுகிறது. இடைநிலையைக் கொண்டு கண்டுபிடிக்கும் சராசரி விலக்கம் மிகக் குறைவாக இருக்கும். சராசரி விலக்கத்தின் தராதர அளவு (coefficient of mean deviation) அல்லது சராசரி விலக்கக்கெழு கீழக்காணும்படி கண்டுபிடிக்கப்படுகிறது.

(1) $\dfrac{\text{கூட்டுச் சராசரியின் மூலம் கிடைத்த சராசரி விலக்கம்}}{\text{கூட்டுச் சராசரி}}$

அல்லது

(2) $\dfrac{\text{இடைநிலை மூலம் கிடைத்த சராசரி விலக்கம்}}{\text{இடைநிலை}}$

கால்மான விலக்கம் (Quartile deviation or semi-interquartile range)

கா.வி (கால்மான விலக்கம்) $= Q = \dfrac{Q_3 - Q_1}{2}$. Q_1 ம் Q_3ம் முதல் மற்றும் மூன்றாம் (absolute measure) கால்மானங்கள். எனவே, அலகுகள் (unit of measurement : கிலோ, லிட்டர், மீட்டர்) வேறுபடும்போது, கால்மான விலக்கங்களைக் கொண்டு ஒப்பிடலாகாது.

கால்மான விலக்கத்தின் கெழு (coefficient of quartile deviation or quartile coefficient of dispersion) $= \dfrac{Q_3 - Q_1}{Q_3 + Q_1}$. இதில் Q_3, Q_1 முறையே மேல்கால்மானம் மற்றும் கீழ்கால்மானம் ஆகும்.

கால்மான விலக்கம் மற்றும் கால்மான விலக்கக்கெழு காண, முதலில் முதல் (கீழ்) மற்றும் மூன்றாம் (மேல்) கால்மானங்களைக் கண்டுபிடித்துக் கொள்ள வேண்டும்.

திட்டவிலக்கம் (Standard deviation = σ or s)

மழையின் அளவுக்கும் அது எந்த அளவுக்கு விவசாயத்தைப் பாதிக்கும் என்பதற்கும் உள்ள தொடர்பினை விளக்க கூட்டுச்சராசரியைவிட, திட்டவிலக்கம் நல்ல முறையில் பயன்படும். உதாரணத்திற்கு, தமிழ்நாட்டில் இராமநாதபுரம் போன்ற மாவட்டங்களின் ஆண்டு மழையளவு அதிகமாக இருந்தாலும் மழை சில நாட்களே (அக்டோபர் -

நவம்பர் மாதங்களில்) அதிகமாகப் பொழிந்து விடுவதால் அந்த மழை விவசாயத்திற்கு மிகுந்த பலனளிக்காமல் போய்விடுகிறது. அப்படியில்லாமல், அதே மொத்த அளவு மழை நான்கு - ஐந்து மாதங்கள் கொஞ்சம் கொஞ்சமாகப் பொழிந்தால் விவசாயத்திற்கு இன்னும் அதிக அளவு பலனுள்ளதாக இருக்கும். எனவே, மழை பல நாட்களுக்குப் பரவிப் பொழிவது ஓரிரு நாட்களிலேயே அதிகமாகப் பொழிவதைவிட நல்லதாக அமையும். பல நாட்கள் பரவி மழை பொழியும்போது சிதறல் அதிகமாக இருக்கும். ஓரிரு நாட்களில் மழை பொழியும்போது கூட்டுச்சராசரி அதிகமாகவும் சிதறல் குறைவாகவும் இருக்கும்.

கூட்டுச்சராசரியிலிருந்து ஒவ்வொரு மதிப்புக்கும் உள்ள வேறுபாடுகளை $(X - \overline{X})$ +, - ஆகிய அடையாளங்கள் போட்டு, தனி மதிப்புக்கள் கூட்டுச் சராசரியை விட அதிகமா குறைவா என்று காட்டப்படுகின்றன. அந்த வேறுபாடுகளின் மொத்தம் பூஜ்யமாக இருக்கும். அதைத் தவிர்க்க, திட்டவிலக்கம் கணக்கிடும்போது ஒவ்வொரு வேறுபாட்டிற்கும் $(X - \overline{X})$ வர்க்கம் (square) கண்டுபிடிக்கப்பட்டு அவை கூட்டப் படுகின்றன.

கொடுக்கப்பட்டுள்ள புள்ளி விபரங்களின் கூட்டுச் சராசரிக்கும் ஒவ்வொரு புள்ளி விபரத்திற்கும் உள்ள வேறுபாடுகளின் வர்க்கங்களுடைய (squared deviations) கூட்டுச் சராசரியின் வர்க்கமூலமே (square root) திட்டவிலக்கமாகும். இது σ (sigma:சிக்மா) என்ற கிரேக்க எழுத்தால் அழைக்கப் படுகிறது. கடைசியில் வர்க்கமூலம் கண்டுபிடிக்கா விட்டால் மீதமிருப்பது மாறுபாடு (variance) என அழைக்கப்படுகிறது. அல்லது திட்டவிலக்கத்தின் வர்க்கம் (square) மாறுபாடு (variance) ஆகும்.

திட்டவிலக்கமும் (sigma = σ) மாறுபாடும் (variance) எடுகோள்களைச் சோதனையிடும்போது (Testing of hypotheses) மிகவும் பயனுள்ளதாக இருக்கின்றன. எனவே, இவற்றை நன்கு புரிந்து கொள்வது மிகுந்த பலன்கள் அளிக்கும்.

முழுத்தொகையின் (Population, Universe) திட்டவிலக்கத்தை (standard deviation) σ என்றும், மாதிரிகளின் திட்டவிலக்கங்களை 's' என்றும் குறிப்பது வழக்கமாக உள்ளது.

$$\text{திட்டவிலக்கம் (தி.வி.)} = \sqrt{\frac{\Sigma (X-\bar{X})^2}{n}} = \sqrt{\frac{\Sigma x^2}{n}}$$

இதில் $\bar{X}$ = கூட்டுச் சராசரி

X = திட்டவிலக்கம் காணப்பட வேண்டிய மாறியின் தனிமதிப்பு

x = கூட்டுச்சராசரிக்கும் தனிமதிப்புக்கும் உள்ள வித்தியாசம்

n = மாறியில் உள்ள உறுப்புக்களின் எண்ணிக்கை

குறிப்பு : N என்பது முழுத்தொகையின் அளவையும் (size of population), n என்பது மாதிரியில் உள்ள உறுப்புக்களின் (number of observations) எண்ணிக்கையையும் குறிக்கப் பயன்படுத்தப்படுகின்றன.

மாறியில் காணப்படும் உறுப்புக்கள் அலைவெண்களுடன் காணப்பட்டால்,

$$\text{தி.வி.} = \sqrt{\frac{\Sigma f(X-\bar{X})^2}{n}} \; ; \quad n = \Sigma f.$$

சில சமயம் மேலே உள்ள nக்குப் பதிலாக n-1 பயன்படுத்தப்படுகிறது. n-1ஐப் பயன்படுத்திக் கண்டுபிடிக்கும் மாதிரித் திட்ட விலக்கம் (sample standard deviation = $\hat{s}$),

முழுத்தொகுதியின் திட்டவிலக்கத்தை (standard deviation of the population = σ) கணிக்க நல்ல நம்பிக்கையுள்ள மதிப்பீடாக இருக்கிறது. ஆனாலும் மாதிரியின் அளவு (sample size) *30*க்கும் அதிகமாக இருக்கும்போது, n-1ஐப் பயன்படுத்திக் காணும் திட்ட விலக்கத்திற்கும் nஐப் பயன்படுத்திக் காணும் திட்டவிலக்கத்திற்கும் அதிக வித்தியாசம் இருப்பதில்லை. உதாரணத்திற்கு, '5'க்கும் (n = 5) '4'க்கும் (n-1; 5-1) உள்ள வித்தியாசம் *1000*க்கும் (n = 1000) 999க்கும் (1000-1, n-1) உள்ள வித்தியாசத்தைவிட அதிக முக்கியத்துவம் வாய்ந்தது. இந்த எண்களைக் கொண்டு, அவற்றிற்குள்ள வித்தியாசத்தைத் தெளிவாகக் காட்டலாம்.

$$\frac{5}{5} = 1; \quad \frac{5}{4} = 1.25$$

$$\frac{5}{1000} = 0.005; \quad \frac{5}{999} = 0.005$$

எனவே, 'n' 1000ஆக இருக்கும்போது n அல்லது n-1ஐப் பயன்படுத்தினால் பெரிய வித்தியாசம் இல்லை. எனவே n-1க்குப் பதிலாக nஐப் பயன்படுத்தலாம். ஆனால், n = 5ஆக இருக்கும்போது n-1ஐப் பயன்படுத்துதல்தான் சரியாகும். மேலும், nஐப் பயன்படுத்தி திட்டவிலக்கம் கண்டுபிடிக்கப்பட்டு இருந்து, நம்பிக்கைக்குரிய மதிப்பீடு வேண்டுமானால், nஐப் பயன்படுத்தி கண்டுபிடிக்கப்பட்ட sஉடன் $\sqrt{n/n\text{-}1}$ ப் பெருக்கி நம்பிக்கைக்குரிய மதிப்பீடாகிய n-1 ஐப் பயன்படுத்திக் கண்டுபிடிக்கப்படும் திட்ட விலக்கத்தினைப் பெறலாம். இதனால், பலசமயம் திட்ட விலக்கத்திற்கான சூத்திரங்களில் n மட்டுமே காணப்படுகிறது; n-1 மிகச் சிலசமயமே, ('n' 30க்கும் குறைவாக உள்ளபொழுது மட்டுமே) காணப்படுகிறது.

முழுத்தொகையின் திட்டவிலக்கம், σ என்றும் மாதிரியின் திட்டவிலக்கம் s என்றும் குறிக்கப்படுவதால்

அவைகளின் மாறுபாடுகள் (variances) முறையே σ^2 என்றும் s^2 என்றும் குறிக்கப்படுகின்றன.

முன்னர் கொடுக்கப்பட்டுள்ள சூத்திரங்களைக் கீழ்வருமாறும் கொடுக்கலாம்.

திட்ட விலக்கம் $= \sqrt{\frac{\Sigma X^2}{n} - \left(\frac{\Sigma X}{n}\right)^2}$ *(1)*

திட்ட விலக்கம் $= \sqrt{\frac{\Sigma fX^2}{n} - \left(\frac{\Sigma fX}{n}\right)^2}$ *(2)*

உண்மையான கூட்டுச் சராசரிக்குப் பதிலாக, சில சமயம் வசதிக்காக ஏதேனும் ஒரு எண்ணை உத்தேச கூட்டுச் சராசரியாகப் (assumed mean) பயன்படுத்த வேண்டி வரலாம். உதாரணத்திற்கு $\bar{X} = 8.2137$ என்று இருந்தால் x $(X - \bar{X})$ காண்பது சிரமமாக ஆகலாம். எனவே $\bar{X}$க்குப் *(8.2317)* பதிலாக *8* அல்லது *10* என்று எடுத்துக் கொண்டு (A) பயன்படுத்தலாம். அப்படிச் செய்யும்போது *(1)*, *(2)* எனக் குறிக்கப்பட்டுள்ள திட்டவிலக்கத்தின் சூத்திரங்கள்

$s = \sqrt{\frac{\Sigma d^2}{n} - \left(\frac{\Sigma d}{n}\right)^2}$ என்றும்

$s = \sqrt{\frac{\Sigma fd^2}{n} - \left(\frac{\Sigma fd}{n}\right)^2}$ என்றும்

ஆகும். இங்கு $d = X - A$

கொடுக்கப்பட்டுள்ள புள்ளி விவரங்கள் சம இடைவெளி உள்ள (Class intervals) சில பிரிவுகளாகப் (Classes) பிரிக்கப்பட்டு இருந்தால் அப்போது

திட்டவிலக்கம் $= \sqrt{\frac{\Sigma fu^2}{n} - \left(\frac{\Sigma fu}{n}\right)^2} \times C$

இங்கு × பெருக்கல் குறி.

C = பிரிவு இடைவெளி (class interval)

$U = \frac{X-A}{C}$. இதில் X = மாறியின் மதிப்பு

A = உத்தேசச் சராசரி

C = பிரிவு இடைவெளி.

திட்டவிலக்கத்திற்குரிய தன்மைகள்

கூட்டுச்சராசரிக்குப் பதிலாக, இடைநிலையையோ, முகடையோ அல்லது வேறு ஏதேனும் ஒரு நிலையான எண்ணையோ (constant) பயன்படுத்தி திட்டவிலக்கம் காணலாம். இருப்பினும், கூட்டுச்சராசரியைப் பயன்படுத்தி காணப்படும் திட்டவிலக்கம்தான், மற்ற வழித் திட்ட விலக்கங்களைவிடக் குறைவாக இருக்கும். எனவே, கூட்டுச் சராசரி வழித் திட்டவிலக்கம் தான் துல்லியம் எனக் கருதப்படுகிறது.

இரண்டு தொடர்களுக்கான கூட்டுச் சராசரியும், திட்டவிலக்கமும் தெரிந்திருந்தால் அந்த இரண்டு தொடர்களுக்கும் பொதுவான திட்டவிலக்கம் கண்டுபிடித்து விடலாம். உதாரணத்திற்கு, ஒரு வகுப்பில் உள்ள 100 மாணவர்களுக்கும் 50 மாணவிகளுக்கும் உள்ள மதிப்பெண்களின் கூட்டுச் சராசரி முறையே 60, 70 என்றும் திட்டவிலக்கங்கள் முறையே 40, 30 என்றும் இருந்தால், அந்த வகுப்பின் மொத்த மாணவ மாணவிகளின் கூட்டுச் சராசரியும், திட்ட விலக்கமும் கண்டுபிடித்து விடலாம்.

கூட்டுச் சராசரி $= \frac{(100)(60) + (50)(70)}{100 + 50}$ ஆகும்.

திட்டவிலக்கம் $= \sqrt{\frac{1}{n}\left(n_1 s_1^2 + n_2 s_2^2 + n_1 d_1^2 + n_2 d_2^2\right)}$

இதில் $n = 100 + 50$

$n_1 = 100$ $\quad s_1^2 = 40^2 = 1600$

$n_2 = 50$ $\quad s_2^2 = 30^2 = 900$

$d_1 = (\bar{X}_1 - \bar{X})$ $\quad d_2 = (\bar{X}_2 - \bar{X})$

$d_1 = (60 - \bar{X})$ $\quad d_2 = (70 - \bar{X})$

கூட்டுச் சராசரியையும் திட்ட விலக்கத்தையும் வைத்து ஒரு பரவலில் உள்ள எத்தனை கூறுகள் எவ்வளவுக்குள் இருக்கும் எனக் கூற இயலும். உதாரணத்திற்கு, ஒரு வகுப்பில் உள்ள 100 மாணவர்களின் மதிப்பெண்கள் இயல்நிலைப் பரவலாக (normal distribution) இருக்கிறதென்றும் அவர்களின் மதிப்பெண்களின் கூட்டுச் சராசரி 60 என்றும் திட்டவிலக்கம் 10 என்றும் கொண்டால் அனேகமாக 34 விழுக்காடு (சதவீதம்) மாணவர்கள் 50க்கும் 60க்கும் இடையில் மதிப்பெண் பெற்றிருப்பார்கள். அதுபோல 60க்கு மேல் 70க்குள் 34 மாணவர்கள் மதிப்பெண் பெற்றிருப்பார்கள். அனேகமாக 95.5 சதவீத (அல்லது 96) மாணவர்கள் 40க்கும் 80க்கும் இடையில் மதிப்பெண் பெற்றிருப்பார்கள் என்று கூட கூற இயலும். இவ்வாறாக திட்ட விலக்கத்தின் தன்மை அமைந்துள்ளது.

மாறுபாட்டுக்கெழு (Coefficient of variation)

திட்டவிலக்கம் ஒரு முழுமையான அளவை (absolute measure) ஆகும். எனவே ஒரு பரவல் எந்த அலகில் (உதாரணத்திற்கு மீட்டர், லிட்டர் ...) கொடுக்கப்பட்டுள்ளதோ அதே அலகில்தான் திட்டவிலக்கமும் தரப்படும். இக்காரணத்தினால் இரு வேறு அலகில் கொடுக்கப்பட்டுள்ள இரு பரவல்களின் சிதறல் அளவுகளை திட்டவிலக்கங்களைப் பயன்படுத்தி ஒப்பிட முடியாது. எனவே சிதறலின் தராதர அளவை அல்லது ஒப்பீட்டு அளவையொன்று அவசியமாகிறது.

மேலும், ஒரு தொகுதியின் உறுப்புகள் பெரிய எண்களாக இருந்தால் கூட்டுச் சராசரியும் பெரியதாக இருக்கும்; சிறிய எண்களாக இருந்தால் கூட்டுச் சராசரியும் சிறியதாக இருக்கும். உதாரணத்திற்கு 1, 2, 3க்கு கூட்டுச் சராசரி 2. ஆனால் 1001, 1002, 1003க்கு கூட்டுச் சராசரி 1002. திட்டவிலக்கம் அப்படியல்ல. 1, 2, 3 என்ற தொடரின் திட்டவிலக்கமும் 1001, 1002, 1003 என்ற தொடரின் திட்டவிலக்கமும் சமமாக இருக்கும். ஏனெனில் இரண்டு தொடர்களிலும் சிதறல் ஒரே அளவுதான். ஆனால் கூட்டுச் சராசரி 2ஆக இருக்கும்போது உள்ள வித்தியாம் 1 (ஒன்று). இது கூட்டுச்சராசரியில் ½ ஆகும். ஆனால் கூட்டுச் சராசரி 1002 ஆக இருக்கும்போதும் வித்தியாசம் 1 தான் என்றாலும், அது கூட்டுச் சராசரியை ஒப்பிடும்போது மிகச்சிறிய எண் ஆக உள்ளது $\left(\frac{1}{1002}\right)$. எனவே ஒரு தொடரின் திட்டவிலக்கத்தை அதே தொடரின் கூட்டுச் சராசரியால் வகுத்து விட்டால் பெரிய எண்ணோ அல்லது சிறிய எண்ணோ ஒப்பிடக்கூடிய வகையில் அமைகிறது.

மேற்கூறிய காரணங்களினால் திட்டவிலக்கக் கெழு (திட்டவிலக்கம் ÷ கூட்டுச் சராசரி) திட்டவிலக்கத்தை விடச் சிறந்ததாகக் கருதப்படுகிறது.

திட்டவிலக்கக்கெழுவை சதவீதத்தில் கூறும்போது அது மாறுபாட்டுக்கெழு என்றும் மாறுவிகிதக் கெழு (coefficient of variation) என்றும் அழைக்கப்படுகிறது.

$$\text{மாறுபாட்டுக் கெழு} = \frac{s}{\bar{X}} \times 100$$

s = மாதிரியின் திட்டவிலக்கம்

$\bar{X}$ = மாதிரியின் கூட்டுச் சராசரி

ஒன்றுக்கு மேற்பட்ட குழு விபரங்கள் கிடைத்தால், அவற்றில் எந்தக்குழு உறுதியான, திடம் வாய்ந்த, சீரானதாக இருக்கின்றது என்று அறிவதற்கு இந்த மாறுபாட்டுக்கெழு உதவுகிறது. உதாரணத்திற்கு, அன்பு என்ற நபருடைய நாள் வருமானத்தின் மாறுபாட்டுக்கெழு (உதாரணத்திற்கு $\frac{400}{100} \times 100 = 400$) ஆதி என்பவரின் நாள் வருமானத்தின் மாறுபாட்டுக்கெழுவை (உதாரணத்திற்கு $\frac{4000}{200} \times 100 = 2000$) விடக் குறைவாக இருந்தால் அன்புடைய நாள் வருமானம் நிலையானதாகவும், சீரானதாகவும் இருப்பதாகக் கூறலாம். இங்கு அன்புடைய சராசரி வருமானத்தைவிட (ரூ.100) ஆதியின் சராசரி வருமானம் (ரூ.200) அதிகமாக இருந்தும் அன்பே இங்கு பாராட்டப்படலாம்.

சிதறல் அளவைகளுக்குள் உள்ள தொடர்பு (Empirical relations between measures of dispersion)

ஓரளவுக்குச் சமச்சீரற்ற பரவலுக்கு (moderately skewed distribution)

சராசரி விலக்கம் $= \frac{4}{5}$ (திட்டவிலக்கம்)

கால்மான விலக்கம் $= \frac{2}{3}$ (திட்டவிலக்கம்)

இயல்பான பரவலுக்கு (Normal distribution)

சராசரி விலக்கம் = 0.7979 திட்டவிலக்கம்

கால்மான விலக்கம் = 0.6745 திட்டவிலக்கம்

தரப்படுத்தப்பட்ட மாறி

மாறிகளுக்கும் அவற்றின் கூட்டுச்சராசரிக்கும் உள்ள வித்தியாசங்களை திட்டவிலக்கத்தால் அளந்தால் அது

தரப்படுத்தப்பட்ட மாறி என அழைக்கப்படுகிறது. அதற்கு அலகுகள் இல்லை. இது 'z' எனும் ஆங்கில எழுத்தால் குறிக்கப்படுகிறது.

$$z = \frac{X - \overline{X}}{s}$$

இதில் X = மாறியின் மதிப்பு

$\overline{X}$ = மாறியின் கூட்டுச் சராசரி

s = மாறியின் திட்டவிலக்கம்

உதாரணத்திற்கு, ஒரு மாணவர் புள்ளியியலில் பெற்ற மதிப்பெண்கள் 84. புள்ளியியலில் அவர் படித்த வகுப்பில் உள்ள மாணவர்களின் கூட்டுச் சராசரி 76 மதிப்பெண் ஆகவும் திட்டவிலக்கம் 10 ஆகவும் இருந்தன. இந்த விவரங்கள் பொருளியல் பாடத்திற்கு முறையே 90, 82, 16 ஆக இருந்தன. அப்படியானால், அந்த மாணவர் எந்தப் பாடத்தில் உண்மையிலேயே அதிக மதிப்பெண்கள் பெற்றுள்ளார் என்று அறிய தரப்படுத்தப்பட்ட மதிப்பைக் காண்பது அவசியமாகிறது.

புள்ளியியலில் அவருடைய தரப்படுத்தப்பட்ட மதிப்பெண் $z = \frac{84-76}{10} = 0.8$

பொருளியலில் அவருடைய தரப்படுத்தப்பட்ட மதிப்பெண் $z = \frac{90-82}{16} = 0.5$

எனவே அந்த மாணவரின் திறமை புள்ளியியலில்தான் (பொருளியலைவிட) அதிகம் என்று கூறலாம். இந்த முடிவு மற்ற மாணவர்களின் திறமையையும் கவனத்தில் கொண்டு எடுத்த முடிவாகும். பொருளியலில் எல்லா மாணவர்களும் அதிக மதிப்பெண்கள் ($\overline{X}$) பெற்றுள்ள காரணத்தினாலும்

அந்தப் பாட மதிப்பெண்களில் மாணவர்களுக்கு இடையே அதிக வேறுபாடுகள் (s) இருந்ததாலும் இந்த குறிப்பிட்ட மாணவரும் அதிக மதிப்பெண்கள் பெற்றுள்ளார். எனவே, அவர் புள்ளியியலில் பெற்ற மதிப்பெண்தான் அவரின் சிறப்பைக் காட்டுகிறது.

இதுபோல தமிழ்நாட்டில் மாநிலப் பாடத்திட்டத்தின் (State Board) மூலம் தேர்வு எழுதுபவர்கள் பெறும் மதிப்பெண்களை மத்திய பாடத்திட்டத்தின் (Central Board) மூலம் தேர்வு எழுதுபவர்கள் பெறும் மதிப்பெண்களுடன் ஒப்பிடுவதில் சிரமம் ஏற்படுகிறது. இந்த இருவேறு முறைகளின் மூலம் இருவர் பெற்ற மதிப்பெண்கள் சமமாக இருந்தபோதும் அம்மாணவர்களைச் சமமான திறன் படைத்தவர்கள் என்று சொல்லமுடிவதில்லை. இந்நிலையில் அவர்கள் உயர்கல்விக்கு இடம்வேண்டி விண்ணப்பித்தால் யாருக்கு முன்னுரிமை கொடுப்பது என்று முடிவு செய்வது சற்று கடினமாக இருக்கிறது. இந்த மாதிரிச் சூழ்நிலையில், அவர்களின் மதிப்பெண்களைத் தரப்படுத்தினால் அவற்றை ஒப்பிட்டுப்பார்ப்பது சரியாகும். உதாரணத்திற்கு, ஒரு மாணவர் மாநிலப் பாடத்திட்டத்தில் 70 மதிப்பெண்கள் பெற்றுள்ளார் என்றும் அதே ஆண்டில் மாநிலப் பாடத் திட்டத்தில் தேர்வு எழுதிய அனைத்து மாணவர்களின் சராசரி மதிப்பெண் 60 என்றும் திட்டவிலக்கம் 10 என்றும் கொண்டால், அந்த மாணவரின் தரப்படுத்தப்பட்ட மதிப்பெண் 1 [(70-60) ÷ 10] என்றாகிறது. மத்திய பாடத்திட்டத்தில் இவ்விவரங்கள் முறையே 50, 30, 10 என்றால், இந்த மதிப்பெண்களைக் கொண்டு தரப்படுத்தப்பட்ட மதிப்பெண் 2 [(50-30) ÷ 10] என்றாகிறது. எனவே மத்திய பாடத்திட்டத்து மாணவர் குறைவான மதிப்பெண்கள் பெற்றிருந்தபோதும் அவருடைய தரப்படுத்தப்பட்ட மதிப்பெண் மாநிலப்

பாடத்திட்டத்து மாணவரின் தரப்படுத்தப்பட்ட மதிப்பெண்ணைவிடக் கூடுதலாக உள்ளது. இவ்வாறாக இருவருடைய தரப்படுத்தப்பட்ட மதிப்பெண்களையும் ஒப்பிடுவது தரப்படுத்தாத மதிப்பெண்களை ஒப்பிடுவதைவிடச் சிறப்பாக இருக்கும். ஏனெனில், இந்த தரப்படுத்தும் முறை மூலம் தனி ஒரு மாணவரது மதிப்பெண், அவருடைய பாடத்திட்டத்தைச் சார்ந்த அனைத்து மாணவர்களின் மதிப்பெண்களையும், அந்த மதிப்பெண்களிடையேயுள்ள வேறுபாடுகளையும் மனதில் கொண்டு கணக்கிடப்படுகிறது. இந்த மாதிரிச் சூழ்நிலைகளில் தரப்படுத்தப்பட்ட மதிப்புக்கள் கணக்கிடும் முறை பொருத்தமானதாகவும் சரியான பொருளுள்ள (Meaningful)தாகவும் இருக்கும்.

லாரென்ஸ் வளைகோடு (Lorenz Curve)

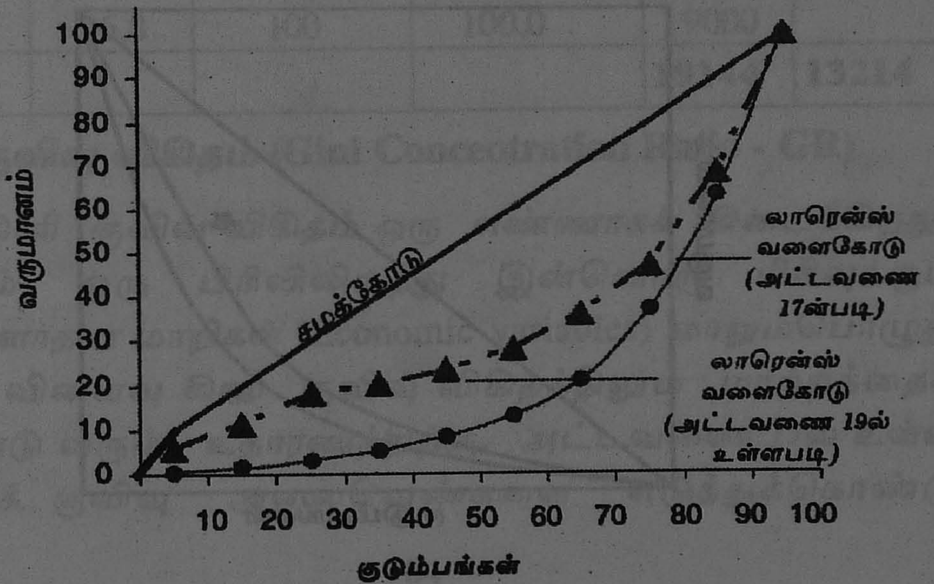
மேக்ஸ் ஓ லாரென்ஸ் (Max O. Lorenz) முதன்முதலில் சொத்துக்கள் தனிமனிதர்களிடையே எவ்வாறு பகிர்ந்தளிக்கப் பட்டுள்ளன என்பதை அளவிட ஒரு வரைபடம் பயன்படுத்தினார். இப்பொழுது பொருளியல் தொடர்பான அனைத்து மாறிகளும் எப்படிப் பரவியுள்ளன என்பதைக் காண்பிக்க லாரென்ஸ் வளைகோடு பயன்படுத்தப்படுகிறது. லாரென்ஸ் வளைகோட்டை வரைய அலைவெண்கள் மற்றும் கொடுக்கப்பட்டுள்ள மாறியின் மதிப்புக்கள் ஆகிய இரண்டிற்குமே சதவீதக் குவிவு மதிப்புக்கள் காணப்படல் வேண்டும். உதாரணத்திற்கு, அட்டவணை 17இல் உள்ள புள்ளி விவரங்களைக் காணலாம்.

அட்டவணை - 17
மாத வருமான விவரங்கள்

வீடுகளின் எண்ணிக்கை	வருமானம் (ரூ) வீடு ஒன்றுக்கு	வீடுகளின் வருமானம் (ரூ)	வீடுகள் சதவீதத்தில்	வீடுகளின் சதவீதக் குவிவு	வருமானம் சதவீதத்தில்	வருமான சதவீதக் குவிவு
10	1000	10000	10	10	0.5	0.5
10	2000	20000	10	20	1.1	1.6
10	3000	30000	10	30	1.6	3.2
10	4000	40000	10	40	2.1	5.3
10	6000	60000	10	50	3.2	8.5
10	9000	90000	10	60	4.7	13.2
10	15000	150000	10	70	7.9	21.1
10	30000	300000	10	80	15.8	36.9
10	50000	500000	10	90	26.3	63.2
10	70000	700000	10	100	36.8	100.0
100		**1900000**				

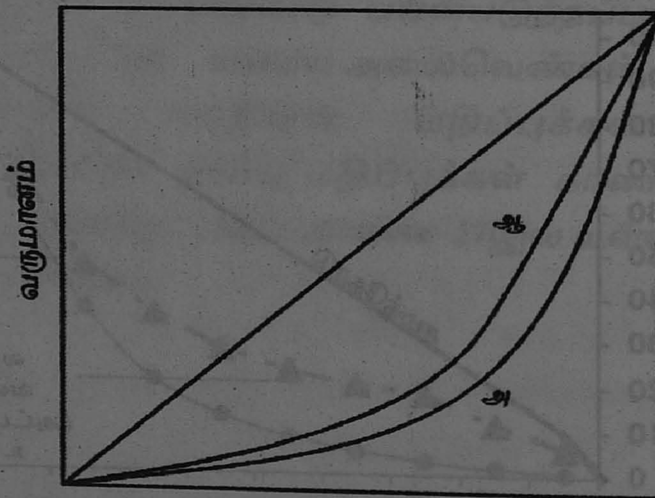
அட்டவணை 17இல் கொடுக்கப்பட்டுள்ள வீடுகளின் சதவீதக்குவிவை கிடைமட்ட (X) அச்சிலும் வருமானச் சதவீதக் குவிவை செங்குத்து (Y) அச்சிலும் இட்டு வரையப்படும் வளைகோடு லாரென்ஸ் வளைகோடாகும்.

வரைபடம் - 15
லாரென்ஸ் வளைகோடு



45° கோணத்தில் வரையப்பட்டுள்ள சமக்கோடு பொருளாதார மாறி (நிலம், வருமானம், செலவு போன்றவை) சமமாகப் பகிர்ந்தளிக்கப்பட்டிருந்தால் எப்படி இருக்கும் என்பதைக் காட்டுகிறது. முதல் 10 சதவீத வீடுகள் 10 சதவீத நிலத்தினையும் கொண்டு அதுபோல ஒவ்வொரு 10 சதவீத வீடுகளும் சமமாக நிலத்தினைப் பெற்றிருந்தால் சமக்கோடுபோல் லாரென்ஸ் கோடும் வந்திருக்கும். ஆனால் பொதுவாக பொருளாதார மாறிகள் அவ்வாறு சமமாகப் பகிர்ந்தளிக்கப்பட்டிருப்பதில்லை. சமமின்மை அதிகரிக்க அதிகரிக்க லாரென்ஸ் வளைகோடு சமக்கோட்டைவிட்டு அதிகமாக தூரமாகிக் கொண்டும் விலகிக் கொண்டும் போகும்; சமக்கோட்டுக்கும் லாரென்ஸ் வளைகோட்டுக்கும் இடையில் உள்ள பகுதியும் பெரிதாகிக் கொண்டிருக்கும். இந்தப் பகுதியின் பரப்பை வைத்தே சமச்சீரின்மை எந்தளவுக்கு உள்ளது என அறியலாம். உதாரணத்திற்கு, லாரன்ஸ் வளைகோட்டுக்கும் சமக்கோட்டுக்கும் இடையில் உள்ள தூரமும் பரப்பளவும் கிராமம் 'அ'க்கு , கிராமம் 'ஆ'க்கு இருப்பதை விட, அதிகமாக இருந்தால் கிராமம் 'அ'வில் அதிகமான அளவுக்கு ஏற்றத்தாழ்வு உள்ளது என்று பொருள்.

வரைபடம் - 16



உதாரணத்திற்கு, அட்டவணை 19இல் கொடுக்கப்பட்டுள்ள மேல்மட்ட வர்க்கத்தினரின் மேல் வரிவிதித்து கீழ்மட்டத்தில் உள்ள 3 பிரிவினருக்குக் கொடுத்த பிறகு குவிவு அலைவெண்களுக்கு லாரென்ஸ் வளைகோடு வரைந்தால் அது வரைபடம் 15ல் உள்ள ←← லாரென்ஸ் வளைகோடாக வந்து வரிவிதிப்பிற்குப் பின்னர் வருமான ஏற்றத்தாழ்வு குறைந்திருப்பதைக் காட்டும்.

அட்டவணை - 18

% P	% Q	P வீடுகளின் சதவீதக் குவிவு அலைவெண்	Q வருமான சதவீதக் குவிவு அலைவெண்	$P_{i-1}Q_i$	$Q_{i-1}P_i$
10	0.5	10	0.5	-	10
10	1.1	20	1.6	16	48
10	1.6	30	3.2	64	128
10	2.1	40	5.3	159	265
10	3.2	50	8.5	340	510
10	4.7	60	13.2	660	924
10	7.9	70	21.1	1266	1688
10	15.8	80	36.9	2583	3321
10	26.3	90	63.2	5056	6320
10	36.8	100	100.0	9000	
				19144	**13214**

கினி குவிவு விகிதம் (Gini Concentration Ratio - GR)

கினி குவிவு விகிதம் ஒரு எண்ணாகக் கிடைக்கிறது. மேலும், ஒரு பிரிவிலிருந்து இன்னொரு பிரிவுக்குப் பொருளாதார மாறிகள் (Economic variables) மாறும்பொழுது அதன் விளைவு கினி குவிவு விகிதத்திலும் மாற்றத்தைக் கொண்டு வரும். உதாரணத்திற்கு, அட்டவணை 17ல் உள்ள சதவீதக் குவிவு அலைவெண்களை எடுத்துக்கொண்டு

அதிலிருந்து அட்டவணை 18ஐ உருவாக்கி கினி குவிவு விகிதம் காணலாம்.

$$GR = \frac{\Sigma P_{i-1}Q_i - \Sigma Q_{i-1}P_i}{10000}$$

$$= \frac{19144 - 13214}{10000} = 0.5930$$

இப்பொழுது கடைசி 3 பிரிவுகளில் (செல்வர்களிடம்) உள்ள 5% வருமானத்தை எடுத்து முதல் 3 பிரிவுகளுக்குக் (வருமானம் குறைவாக உள்ளவர்களுக்கு) கொடுப்பதாகக் கொள்வோம். இப்போது அந்த ஊரில் வருமான ஏற்றத்தாழ்வு குறைய வேண்டும்; அதை கினி குவிவு விகிதம் குறைந்து காட்ட வேண்டும். மேலே சொன்ன மாற்றத்தால் அட்டவணை 18, அட்டவணை 19 ஆக மாறும்.

அட்டவணை - 19

% P	% Q	P வீடுகளின் சதவீதக் குவிவு அலைவெண்	Q வருமான சதவீதக் குவிவு அலைவெண்	$P_{i-1}Q_i$	$Q_{i-1}P_i$
10	5.5	10	5.5	-	110
10	6.1	20	11.6	116	348
10	6.6	30	18.2	364	728
10	2.1	40	20.3	609	1015
10	3.2	50	23.5	940	1410
10	4.7	60	28.2	1410	1974
10	7.9	70	36.1	2166	2888
10	10.8	80	46.9	3283	4221
10	21.3	90	68.2	5456	6820
10	31.8	100	100.0	9000	
				23344	19514

$$GR = \frac{\Sigma P_{i-1}Q_i - \Sigma Q_{i-1}P_i}{10000} = 0.3830$$

இவ்வாறாக வருமானத்தினை அதிக வருமானம் உள்ளவர்களிடமிருந்து குறைவான வருமானம் உள்ளவர்களுக்கு மாற்றும்போது, வருமான ஏற்றத்தாழ்வு குறைகிறது; அதனை கினி குவிவு விகிதமும் காட்டுகின்றது.

'சென்'னின் குறியீடு (Sen's Index)

அமர்த்யா குமார் சென் (Amartya Kumar Sen) என்பவர் வறுமை பற்றிய ஆய்வில் தலை எண்ணிக்கை விகிதம் (Head Count Ratio) சில உண்மைகளைக் காட்ட இயலாமல் இருக்கிறது என்றும், அதனைச் சரி செய்ய வருமான இடைவெளி விகிதம் (Income Gap Ratio) பொருத்தமாக இருக்கும் என்றும் தன் கருத்தினை முன்வைத்தார். உதாரணத்திற்கு, இரண்டு ஊர்களில் தலை எண்ணிக்கை விகிதப்படி ஒரே அளவுக்கு வறுமையின் அளவு உள்ளது என்றாலும், அவ்விரு ஊர்களிலும் வறுமையின் பாதிப்பு (incidence of poverty) வேறு விதமாக இருக்க வாய்ப்புண்டு என்பதை சென் விளக்கினார். உதாரணத்திற்கு, மாதாந்திர தலைவீத நுகர்வுச் செலவு (Monthly Percapita Consumption Expenditure = MPCE) ரூ. 450ஐ வறுமைக்கோடு என்றும் இரண்டு ஊர்களிலும் 50 சதவீதத்தினர் வறுமைக்கோட்டிற்குக் கீழே உள்ளனர் என்றும் கொள்ளலாம். அப்படியெனில் வறுமையின் தாக்கம் அவ்விரு ஊர்களிலும் சமமாக இருக்க வேண்டும் என்பதில்லை. ஓர் ஊரில் அது அதிகமாகவும், மற்றோர் ஊரில் அது குறைவாகவும் இருக்கலாம். வறுமையின் தாக்கம் வருமானம் வறுமைக்கோட்டிலிருந்து எவ்வளவு தூரம் விலகி இருக்கின்றது என்பதைப் பொறுத்து அமையும். உதாரணத்திற்கு, ஓர் ஊரில் வறுமைக்கோட்டிற்குக் கீழேயுள்ள ஒவ்வொருவரும் ரூ.430ஐ மாதாந்திர நுகர்வுக்காகச் செலவு

செய்கின்றார்கள் என்றும், மற்றோர் ஊரில் வறுமைக்கோட்டிற்குக் கீழே உள்ளவர்கள் மிகவும் விலகி ரூ.300 மட்டுமே மாதாந்திர நுகர்வுக்காகச் செலவு செய்கின்றார்கள் என்றும் கொண்டால் வறுமையின் சுமை இரண்டாவது ஊரில் அதிகம் எனக் கொள்ளலாம். எனவே, வறுமைக்கோட்டிற்கும் உண்மையான வருமானத்திற்கும் அல்லது செலவுக்கும் உள்ள வேறுபாட்டினை அளப்பது அவசியம் என்பதும் அதைக்கொண்டு வறுமையை அளக்கும் குறியீடு அமைய வேண்டும் என்பதும் 'சென்'னின் கருத்து.

சென் முதலில் வறுமைக்கோட்டுக்கும் கீழ் உள்ள ஒவ்வொருவரின் வறுமை இடைவெளியை அளவிட வேண்டும் என்கிறார்.

வறுமை இடைவெளி (Poverty Gap) = $g_i = \pi - y_i$

இதில், π என்பது வறுமைக்கோட்டைக் குறிக்கும் வருமானம்

y_i என்பது வறுமைக்கோட்டுக்கும் கீழ் உள்ளவர்கள் ஒவ்வொருவரின் வருமானம்

எனவே, மொத்த வறுமை இடைவெளி = $g = \Sigma g_i$

மற்ற இரு வறுமை அளவீடுகள்

(1) தலை - எண்ணிக்கை விகிதம் (Head Count Ratio) = q/n
(q = வறியவர்களின் எண்ணிக்கை)

(2) வருமான இடைவெளி விகிதம் = $I = g/q\pi$
(Income Gap Ratio)

ஏழைகளின் சராசரி வருமானம் = $y^* = \Sigma y / q$

வறுமை இடைவெளியின் சராசரி $g^* = \pi - y^* = g/q$

இதிலிருந்து வருமான இடைவெளி விகிதம் (I) = g*/q என்றும் சொல்லலாம்.

இவைகளைப் பயன்படுத்தி சென்னின் வறுமை அளவையாக P = H {I + (1-I)G} பெறப்படுகிறது. இதில்

G என்பது கினியின் (GINI) வருமானக் குவிவு அளவை.

I என்பது வருமான இடைவெளி விகிதம்

H என்பது வறியவர்களின் தலை எண்ணிக்கை விகிதம்.

4.விலக்கம், கோட்டம் & தட்டைத்தன்மை (MOMENT, SKEWNESS AND KURTOSIS)

விலக்கம் (Moment)

விலக்கத்தினை முதல், இரண்டாம், மூன்றாம் ... என அளவிடலாம்.

முதல் விலக்கம் ஒரு மாறியின் கூட்டுச் சராசரியாகும். உதாரணத்திற்கு 2, 3, 7, 8, 10 ஆகிய மதிப்புக்களின் முதல் விலக்கம் $\frac{2+3+7+8+10}{5} = \frac{30}{5} = 6$ இதுவே கூட்டுச் சராசரியாகும்.

அந்த மதிப்புக்களின் இரண்டாவது விலக்கம் $\left(\frac{\Sigma X^2}{n}\right) = \overline{X}^2$

$$= \frac{2^2+3^2+7^2+8^2+10^2}{5} = \frac{226}{5} = 45.2$$

அந்த மதிப்புக்களின் மூன்றாவது விலக்கம் $= \frac{\Sigma X^3}{n} = \overline{X}^3$

$$= \frac{2^3+3^3+7^3+8^3+10^3}{5} = \frac{1890}{5} = 378$$

இவ்வாறாக மற்ற விலக்கங்களையும் காணலாம். அதுபோல, முதல், இரண்டாவது, மூன்றாவது, ..., விலக்கங்களை அதனுடைய கூட்டுச் சராசரியினடிப்படையில் (moments about the mean) பெறலாம். கீழே அவை கொடுக்கப்பட்டுள்ளன.

$$m_1 = \frac{\Sigma (X-\overline{X})}{n} = \frac{(2-6)+(3-6)+(7-6)+(8-6)+(10-6)}{5} = \frac{0}{5} = 0$$

கூட்டுச் சராசரியினடிப்படையிலான முதல் விலக்கம் 0 ஆக இருக்கும். இரண்டாவது விலக்கம் மாறுபாடாக (variance) அமையும்.

$$m_2 = \frac{\Sigma (X-\bar{X})^2}{n}$$

மூன்றாவது விலக்கம் : $m_3 = \frac{\Sigma (X-\bar{X})^3}{n}$

கூட்டுச் சராசரிக்குப் பதிலாக வேறு ஏதேனும் ஒரு எண்ணின் அடிப்படையில் கூட பலவகை விலக்கங்களைப் பெற முடியும். உதாரணத்திற்கு எண் 4ன் அடிப்படையிலான

முதல் விலக்கம்

$$m'_1 = \frac{\Sigma (X-4)}{n} = \frac{(2-4)+(3-4)+(7-4)+(8-4)+(10-4)}{5}$$

$$= 2$$

இரண்டாம் விலக்கம்

$$m'_2 = \frac{\Sigma (X-4)^2}{n} = \frac{(2-4)^2+(3-4)^2+(7-4)^2+(8-4)^2+(10-4)^2}{5}$$

$$= \frac{66}{5} = 13.2$$

இவ்வாறாக மற்ற விலக்கங்களையும் பெறலாம். அதுபோலவே அலைவெண் பரவல்களுக்கும் விலக்கங்கள் காணலாம்.

கோட்டம் (SKEWNESS)

ஒரு மாறி (variable)யின் குணங்கள் பற்றித் தெரிய விரும்பும்போது அதன் மையப்போக்கு எப்படியுள்ளது என்பது பற்றியும் அந்த மாறி எந்த அளவுக்குப் பரவி அதன் மையப்போக்கினை விட்டு விலகியுள்ளது என்பது பற்றியும் அறிந்து கொள்வது எப்படி என்று இதுவரை விவரிக்கப்பட்டது. கோட்டம் பற்றி அறிந்து கொள்வதன் மூலம் அந்த மாறி எப்படிப் பரவியுள்ளது என்று தெரிந்து கொள்ளலாம். அதாவது, மாறி அதன் மைய அளவுகளின் இருபுறமும் சமமாகப் பரவியுள்ளதா, அல்லது ஏதேனும் ஒருபுறம் மட்டும்

அதிகமாகப் பரவியுள்ளதா, அப்படி ஒரு புறம் மட்டும் அதிகமாகப் பரவியிருந்தால், அதன் மைய அளவின் இடதுபுறமா அல்லது வலதுபுறமா என்று அறிவது பலசமயம் பயனுள்ளதாக இருக்கும். முதலிலேயே கூறியபடி ஒவ்வொரு மாறியும் வெவ்வேறு வகையானப் பரவலைக் கொண்டிருக்கலாம். பொதுவாக இயற்கையால் தீர்மானிக்கப் படுகின்ற போக்கைக் கொண்ட மாறிகள் இயல்நிலைப் பரவலாகச் சமச்சீராக பரவியிருப்பதாக அறிகிறோம். மனிதனுடைய கொள்கைகளாலும் உற்பத்தி உறவுகளாலும் சமூகப் பொருளாதார அரசியல் அதிகார ஆதிக்கங்களாலும் நிர்ணயிக்கப்படுகின்ற மாறிகள் சமச்சீரற்ற (இடதுபுறம் அல்லது வலதுபுறம் குவிக்கப்பட்டு) பரவல்களைக் கொண்டுள்ளன.

அட்டவணை - 20
சமச்சீர் பரவலுக்கான எடுத்துக்காட்டு

மாறியின் மதிப்பு (X)	அலைவெண் (f)	fx	குவிவு அலைவெண் cf
1	6	6	6
2	9	18	15
3	10	30	25
4	11	44	36
5	10	50	46
6	9	54	55
7	6	42	61
மொத்தம்	61	244	

கூட்டுச் சராசரி $\overline{X} = \frac{\Sigma fx}{\Sigma f} = \frac{244}{61} = 4$; Mode = முகடு = 4

இடைநிலை $= \frac{61+1}{2} = \frac{62}{2} = 31$ (அதாவது 31ஆவது X)

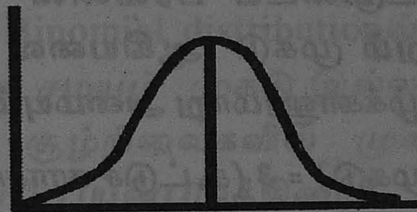
எனவே 4.

இவ்வாறாக சமச்சீரான பரவல் கூட்டுச் சராசரி, முகடு மற்றும் இடைநிலை ஆகியவற்றைச் சமமாகக் கொண்டு இருக்கும். அப்படியிருந்தால் அந்த மாறி அதன் மையப்போக்கு அளவைகளின் இருபுறமும் சமமாகப் பரவியிருக்கும்.

இப்படிப்பட்ட சமன்சீர்பெற்ற பரவலில் இடைநிலை சரியாக மேல்கால்மானத்திற்கும் கீழ்கால்மானத்திற்கும் இடையில் அமைந்திருக்கும். அதாவது (Q_3 - M = M - Q_1)

சமச்சீர் அலைவெண் வளைகோடு வரைந்தால் அது வரைபடம் 16(அ) போல் இருக்கும். இந்தப் படத்தினை செங்குத்துக் கோட்டில் வைத்து இரண்டாக மடித்தால் இருபுறமும் உள்ள வளைகோடுகள் ஒன்றையொன்று தொட்டுக் கொண்டு விலகாமல் இருக்கும்.

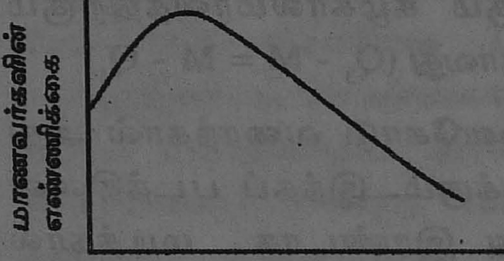
வரைபடம் - 16அ



கொடுக்கப்பட்டுள்ள பரவலானது சமச்சீர் பரவலாக இல்லாதபோது அப்பரவல் கோட்டம் உடையது என்று கூறப்படுகிறது. கோட்டம் உடைய பரவல்கள் நேர்கோட்டம் உடையவை எனவும் எதிர்கோட்டம் உடையவை எனவும் பிரிக்கப்படுகின்றன. ஏற்கனவே வரைபடங்கள் 8 மற்றும் 9களில் காட்டியதுபோல ஒரு போட்டித் தேர்வில் மாணவர்கள் பெற்ற மதிப்பெண்கள் பரவலும் (படம் 17அ) ஒரு பல்கலைக்கழகத் தேர்வில் மாணவர்கள் பெற்ற மதிப்பெண்கள் பரவலும் (படம் 17 ஆ) முறையே நேர்கோட்டமுடையதாகவும் [வலதுபுறம் வால் (tail) நீண்டு] எதிர்கோட்டம் (இடதுபுறம்

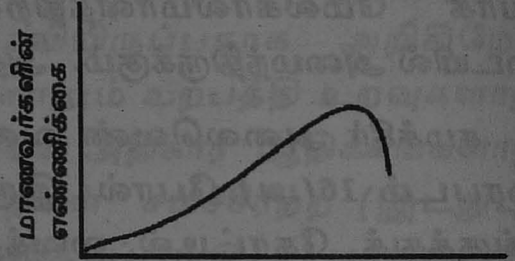
வால் நீண்டு) உடையதாகவும் இருக்கும். இம்மாதிரிப் பட்ட கோட்டமுடைய பரவல்களின் கூட்டுச் சரர்சரிகள் சமமாக இருக்கலாம்; அதுபோலவே அவற்றின் திட்ட விலக்கங்களும் சமமாக இருக்கலாம்.

வரைபடம் - 17



ஒரு போட்டித் தேர்வில் மாணவர்கள் பெற்ற மதிப்பெண்கள்

17. அ



ஒரு பல்கலைக்கழகத் தேர்வில் மாணவர்கள் பெற்ற மதிப்பெண்கள்

17. ஆ

ஒரு கோட்டமுடைய பரவலின் கூட்டுச் சராசரி, இடைநிலை மற்றும் முகடு ஆகியவை சமமாக இராது. அவற்றின் உறவு கீழ்க்காணுமாறு அமையும்.

(கூட்டுச் சராசரி - முகடு) = 3 (கூட்டுச் சராசரி - இடைநிலை)

கோட்ட அளவைகள்

கோட்டமுள்ள பரவலில் முகடு எந்தப் பக்கம் உள்ளதோ அந்தப் பக்கமே அதன் கூட்டுச் சராசரியும் இருக்கும். இவ்வாறாக முகட்டுக்கும் கூட்டுச் சராசரிக்கும் உள்ள வித்தியாசமே கோட்டத்தை அளக்க உதவுகிறது. இந்த வித்தியாசத்திற்கும் கோட்டத்தின் அளவுக்கும் நேரிடை உறவு உள்ளது.

கோட்ட அளவை = (கூட்டுச் சராசரி - முகடு) இதில் +குறி வந்தால் (கூட்டுச் சராசரி > முகடு), அப்பரவல் நேர் கோட்டமுடையது என்றும், -குறி வந்தால் (கூட்டுச் சராசரி < முகடு) அப்பரவல் எதிர்கோட்டம் உடையது என்றும் ஆகும்.

மேலே கூறப்பட்டுள்ள கோட்ட அளவை, கூட்டுச் சராசரியும் முகடும் எந்த அலகில் (Unit of measurement) உள்ளனவோ, அதே அலகிலேயே அமையும். எனவே, இரு வேறுபட்ட அலகில் உள்ள (உதாரணத்திற்கு கிலோ கிராமிலும், மீட்டரிலும்) பரவல்களின் கோட்டத்தை ஒப்பிட முடியாது. இதற்காக, கோட்டத்தின் தராதர அளவை பயன்படுகிறது. இதனை கார்ல் பியர்சன் (KARL PEARSON) தந்ததால், இது கார்ல் பியர்சனின் கோட்டக்கெழு (Karl Pearson's Coefficient of skewness) என அழைக்கப்படுகிறது.

$$\text{கா.பி.கோ.கெ.} = \frac{\text{கூட்டுச் சராசரி - முகடு}}{\text{திட்டவிலக்கம்}}$$

$$= \frac{\overline{X} - \text{Mode}}{s}$$

ஒரு மாறியின் முகட்டினைக் காண்பதில் சில சமயம் சிரமங்கள் இருக்கலாம்; சில சமயம் ஒன்றிற்கு மேற்பட்ட (Bimodal - இது Binomial distributionல் வரலாம்) முகடுகள் இருக்கலாம்; சில சமயம் முகடு இல்லாமலே இருக்கலாம். இப்படிப்பட்ட சூழ்நிலைகளில் முகட்டுக்குப் பதிலாக இடைநிலையைப் பயன்படுத்தினார். கூட்டுச் சராசரிக்கும் முகட்டுக்கும் உள்ள வித்தியாசம் கூட்டுச் சராசரிக்கும் இடைநிலைக்கும் உள்ள வித்தியாசத்தைப் போல் மூன்று மடங்காக இருக்குமாதலால், முன்னதற்குப் பதில் பின்னதைப் பயன்படுத்தி

$$\text{கோட்டக்கெழு} = \frac{3(\overline{X} - \text{Med})}{s}$$

$$= \frac{3\ (\text{கூட்டுச் சராசரி - இடைநிலை})}{\text{திட்டவிலக்கம்}}$$

என்று கார்ல் பியர்சன் விளக்கினார். இவையிரண்டும் முறையே பியர்சனின் முதல் மற்றும் இரண்டாம் கோட்டக்கெழு என

அழைக்கப்படுகின்றன. கோட்டக் கெழுவின் மதிப்பானது நேர்மறை (+)யாகவோ, எதிர்மறை (-)யாகவோ பூஜ்ய (0)மாகவோ இருக்கும்.

$\left(\frac{\overline{X} - Mode}{s}\right)$ -1க்கும் +1க்கும் இடையில் இருக்கும் எனச் சொல்லப்படுகிறது. அதனால் $\frac{3(\overline{X} - Med)}{s}$ -3க்கும் +3க்கும் இடையில் இருக்கலாம். கோட்டக்கெழு பூஜ்யமானால், அங்கு சமச்சீர் பரவல் உள்ளது எனப் பொருள்.

கால்மானங்களைக் கொண்டும் கோட்ட அளவைகளை நிர்ணயிக்கலாம்.

Q_1ஐக் கீழ்கால்மானம் என்றும்

Q_2ஐ மையக் கால்மானம், இரண்டாம் கால்மானம் அல்லது இடைநிலை என்றும்

Q_3ஐ மேல்கால்மானம் என்றும் கொண்டு

கால்மானக் கோட்டக்கெழு நிர்ணயிக்கப்படுகிறது.

கால்மானக் கோட்டக்கெழு

(Quartile coefficient of skewness) $= \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{(Q_3 - Q_2) + (Q_2 - Q_1)}$

இதனை $\frac{Q_3 - 2Q_2 + Q_1}{Q_3 - Q_1}$ எனவும் கொண்டு வரலாம்.

மேலே கூறியது போல, சதவீதமானங்களைக் கொண்டும் கோட்டக் கெழுவினை நிர்ணயிக்கலாம்.

10-90 சதவீதக் கோட்டக்கெழு

(10-90 percentile coefficient of skewness) $= \frac{(P_{90} - P_{50}) - (P_{50} - P_{10})}{(P_{90} - P_{50}) + (P_{50} - P_{10})}$

$= \frac{P_{90} - 2P_{50} + P_{10}}{P_{90} - P_{10}}$

விலக்கக் கோட்டக்கெழு

(Moment coefficient of skewness) $= a_3 = \frac{m_3}{s^3} = \frac{m_3}{\sqrt{m_2}^3}$

இதில், $m_3 = \frac{\Sigma (X-\bar{X})^3}{n}$

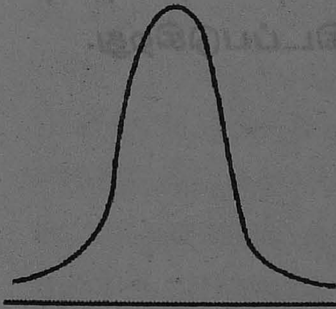
$$m_2 = \frac{\Sigma (X-\bar{X})^2}{n}$$

தட்டைத்தன்மை

கொடுக்கப்பட்டுள்ள மாறி (variable)யின் அலைவெண் வளைகோடு தட்டையான உச்சியையோ கூர்மையான உச்சியையோ கொண்டிருக்கும். இயல்நிலை வளைகோட்டின் உச்சியைவிட ஓர் அலைவெண்ணின் வளைகோடு தட்டையாக உள்ளதா அல்லது கூர்மையாக உள்ளதா என்று கணிப்பது தட்டைத்தன்மையின் அளவை (Measure of Kurtosis) ஆகும்.

வரைபடம் - 18

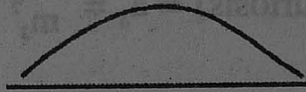
தட்டைத்தன்மைகள்



LEPTOKURTIC

கூர்மையான உச்சி

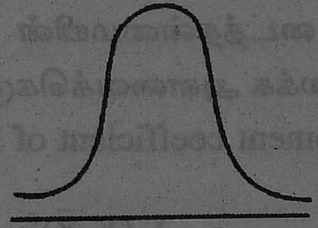
(அ)



PLATYKURTIC

தட்டையான உச்சி

(ஆ)



MESOKURTIC

இயல்நிலை உச்சி (அல்லது) சாதாரண உச்சி (இ)

இத்தட்டைத்தன்மை ஒரு மாறியின் அலைவெண்கள் எவ்வாறு வேறுபடுகின்றன என்பதைப் பொறுத்து அமையும். பரவலின் மையப்பகுதியில் அலைவெண்கள் திடீரென அதிகமாகி

விட்டால் அப்போது வரைபடம் 18இல் முதல் படம் போல் காட்சியளிக்கும். அலைவெண்கள் மெதுவாகக் கூடி, மையப் பகுதியில் ஓரளவு சமமாக இருந்து பின்னர் குறைந்தால் அந்த அலைவெண் வளைகோடு படம் 'ஆ'வில் உள்ளது போல் இருக்கும்.

$$\text{தட்டைத்தன்மையின் சதமான அளவை} = \frac{\text{90ஆவது சதமானத்திற்கும் 10ஆவது சதமானத்திற்கும் உள்ள வேறுபாடு}}{\text{கால்மான விலக்கம்}}$$

$$= \frac{P_{90} - P_{10}}{\left(\frac{Q_3 - Q_1}{2}\right)}$$

$$\text{தட்டைத்தன்மையின் சதமான அளவைக்கெழு} = \frac{\left(\frac{Q_3 - Q_1}{2}\right)}{P_{90} - P_{10}}$$

(Percentile coefficient of kurtosis)

இதில் $\frac{(Q_3 - Q_1)}{2}$ இரண்டு கால்மான இடைவெளிகளின் பாதி (semi-interquartile range) ஆகும்.

கூட்டுச் சராசரியின் அடிப்படையிலான விலக்கத்தைக் கொண்டும் தட்டைத்தன்மைக்கெழு கணக்கிடப்படுகிறது.

தட்டைத்தன்மையின்
விலக்க அளவைக்கெழு
(moment coefficient of kurtosis) $= a_4 = \frac{m_4}{m_2^2}$

இதில் $m_4 = \frac{\Sigma (X-\bar{X})^4}{n}$

$$m_2 = \frac{\Sigma (X-\bar{X})^2}{n}$$

முழுமையின் விலக்கம், கோட்டம் மற்றும் தட்டைத்தன்மை

ஒரு முழுமையிலிருந்து (from population or universe) மாதிரி (தொகுதி அல்லது sample) எடுக்கப்பட்டு அந்த

மாதிரிகளுக்கு கணக்கிடப்படும் விலக்கம், கோட்டம் மற்றும் தட்டைத்தன்மை இலத்தின் குறியிலும், அந்த முழுமையின் விலக்கம், கோட்டம் மற்றும் தட்டைத்தன்மை கிரேக்க குறியிலும் குறிக்கப்படுகின்றன. மாதிரியின் விலக்கம் m எனவும், முழுமையின் விலக்கம் μ (mu, Greek letter μ) எனவும் குறிக்கப்படுகின்றன.

அதுபோல, மாதிரியின் கோட்ட அளவை a_3 எனவும் முழுமையின் கோட்ட அளவை α_3 (alpha : Greek letter) எனவும் குறிக்கப்படுகின்றன.

தட்டைத்தன்மை மாதிரிக்கு a_4 எனவும் முழுமைக்கு α_4 (Greek letter alpha) எனவும் குறிக்கப்படுகின்றன.

ஞ

5. ஒட்டுறவு
(CORRELATION)

இதுவரையில் ஒரு மாறியின் (Univariate) தன்மைகளைப் பலவாறாகப் பல கோணங்களிலிருந்து ஆய்வு செய்வது எப்படி என்று விளக்கப்பட்டது. ஒரு மாறியின் மையப்போக்குகள் யாவை? ஒரு மாறி எவ்வாறு பரவி, சிதறி, விலகி, நெருங்கி, குவிந்திருப்பதை ஆய்வு செய்வது? போன்ற கேள்விகளுக்கான செய்திகள் இதுவரையிலும் தரப்பட்டிருந்தன. ஒன்றுக்கும் மேற்பட்ட மாறிகளை அவற்றின் மையப்போக்கு, சிதறல், விலகல், குவிதல் போன்றவற்றைக் கொண்டு எவ்வாறெல்லாம் ஒப்பிடலாம் அல்லது வேறுபடுத்தலாம் என்றும் அறியப்பட்டது.

அன்றாட வாழ்வில் ஒட்டுறவு கொண்டுள்ள பல மாறிகளைக் காணலாம். அனேகமாக எல்லாவித மாறிகளும் குறைந்தது ஏதேனும் ஒன்று அல்லது இரண்டு மாறிகளுடன் தொடர்பு கொண்டுதான் இருக்கும். எந்த மாறியுடனும் தொடர்பில்லாத மாறி என்று குறிப்பிடும்படியாக ஏதேனும் மாறி இருப்பதாகத் தெரியவில்லை. தனியாக இருக்கும் எரிமலையின் குணங்கள் கூட ஏதேனும் மற்ற குணங்களுடன் தொடர்பு கொண்டு இருக்கலாம். எனவே, இரண்டு அல்லது இரண்டுக்கு மேற்பட்ட மாறிகளுக்கிடையே உள்ள உறவு முறைகளை அறிவது எப்படிப்பட்ட ஆய்வுக்கும் அவசியமாகிறது. அனேகமாக, இரண்டு அல்லது இரண்டுக்கும் மேற்பட்ட மாறிகளுக்கிடையேயான உறவுகளை ஆய்வு செய்யாமல், எந்த ஆய்வும் முழுமை பெறாது என்று கூடக் கூறலாம்.

மாறிகளில் பலவகைகள் உள்ளன. எண்களால் கூறக் கூடியது (உயரம், வருமானம், மழையளவு); எண்களால் கூறக்கூடியவற்றில் முழு எண்கள் மட்டுமே (discrete) எடுத்துக் கொள்வது (மனிதர்கள், வீடுகள்); பின்னங்களாகவும், தசம எண்களாகவும் (உயரம், எடை, மதிப்பெண்கள்) வருவன (continuous). சில மாறிகள் சதவீதத்தில் வரும்போது 100க்கு மேல் மதிப்புக்கள் எடுக்கமாட்டா. சில மாறிகளின் மதிப்புகளுக்கு எல்லையே இல்லாமல் இருக்கும் (ஒரு நாளின் மழையளவு). சில மாறிகளின் மதிப்பு -1க்கும் +1க்கும் இடையில் இருக்கலாம்; சில மாறிகள் 0வுக்கும் +1க்கும் இடையில் மட்டும் மதிப்பு பெறலாம். சில மாறிகள் 0 அல்லது 1 மட்டுமே மதிப்பாகப் பெறலாம். இவ்வாறாகப் பலவகையான மாறிகள் இருந்தபோதும், இந்த அத்தியாயத்தில் தொடர் எண்களைப் பெறக்கூடிய மாறிகளைக் (Continuous variables) கொண்டே உறவுகள் கூறப்படுகின்றன. சார்புடைய மாறி (Dependent variable, explained variable, outcome variable) 0வுக்கும் 1க்கும் இடையே மட்டும் மதிப்பு எடுக்கும் என்றால் அச்சமயங்களில் வேறு முறைகளைப் பயன்படுத்த வேண்டும்; அவை இங்கு விளக்கப்படவில்லை.

மாறிகளுக்கிடையேயுள்ள உறவுகள் நேர்கோடுகளில் (linear) அல்லது பலவிதமான வளைகோடுகளில் (non-linear) வரையக்கூடியதாக இருக்கலாம். வளைகோடு உறவுகளிலும் பலவகைகள் உள்ளன (Quadratic, Quartic, Hyperbolic, Parabolic). ஆனாலும் இந்த அத்தியாயத்தில் நேர்கோடு உறவுமுறை மட்டுமே விளக்கப்படுகிறது.

ஒவ்வொரு மாறியும் பல மாறிகளின் மதிப்புக்களைப் பாதிக்கக் கூடியதாகவும், பல மாறிகளின் மதிப்புக்களால் பாதிக்கப்படக்கூடியதாகவும் இருக்கலாம். ஆனாலும், இந்த அத்தியாயத்தில் இரண்டு தொடர்மாறிகள் (Continuous variables) இடையே உள்ள நேர்கோட்டு உறவுகள் (Simple Linear Relationship) மட்டுமே விவரிக்கப்படுகின்றன.

இரண்டு மாறிகளின் மதிப்பும் ஒன்றை ஒன்று சார்ந்திருக்கலாம் (Mutual dependence); இரண்டு மாறிகளுக்கிடையே காரணகாரியத் தொடர்பு அல்லது காரண விளைவுத் தொடர்பு (Cause and effect relationship) இருக்கலாம். அத்தொடர்புகள் நேரடியாகவோ (direct) மறைமுகமாகவோ (indirect) இருக்கலாம். உதாரணமாக மழையின் அளவுக்கும் விளைச்சலுக்கும் உள்ள தொடர்பு நேரடித் தொடர்பாகும். அதேசமயம், மழையின் அளவுக்கும் விளைபொருட்களின் விலைகளுக்கும் உள்ள தொடர்பு மறைமுகத் தொடர்பாகும். மழையின் அளவு நேரடியாக விளைச்சலைப் பாதித்து அதன் மூலம் விலைவாசியையும் பாதிக்கலாம். சில மாறிகளுக்கிடையேயுள்ள உறவுகளை ஆய்ந்து பார்த்தால் எது காரணம் எது விளைவு என்று முடிவு செய்வதும் கடினமாக இருக்கலாம். எந்த மாறி எந்த மாறியைச் சார்ந்துள்ளது என்று தீர்க்கமாகச் சொல்ல முடியாத நிலையும் வரலாம்; அல்லது அவற்றிற்கிடையேயுள்ள காரண காரிய உறவுசூழ்நிலைக்கேற்ப மாறலாம். உதாரணத்திற்கு குடும்பத்தின் செலவு குடும்ப வருமானத்தைப் பொறுத்து அமையுமா? அல்லது குடும்பத்தின் செலவு குடும்ப வருமானத்தை நிர்ணயிக்குமா? சிலர் குடும்ப வருமானத்தை குடும்ப செலவுக்கேற்ப பெருக்கிக் கொள்பவர்களாக இருக்கலாம். சிலர் குடும்ப வருமானத்திற்குள் குடும்பச் செலவை வைத்துக் கொள்பவர்களாக இருக்கலாம். அதுபோல, கல்விக்கும் குடும்ப வருமானத்திற்கும் இடையேயுள்ள உறவின்முறையை (nature of relationship) நிர்ணயிப்பதிலும் சிரமம் இருக்கலாம். வருமானம் அதிகம் உள்ளவர்களால் அதிகம் படிக்க இயலும். அதிகம் படிப்பவர்களால் அதிக வருமானம் ஈட்ட முடியலாம்.

சில சமயங்களில் சில மாறிகளிடையே உள்ள உறவு வெளிப்படையாகத் தெரியாமல் இருக்கலாம். உதாரணத்திற்கு, புகைப்பிடிக்கும் பழக்கத்திற்கும் தலையின் உரோமம்

கறுப்பாக இருப்பதற்கும் ஏதேனும் தொடர்பு உள்ளதா என்று அறிவதற்கு புள்ளிவிபரம் சேகரித்துப் பார்த்தால், புகைப்பிடிப்பவர்களின் தலையில் உரோமம் கறுப்பாகவும், புகைப்பிடிக்காதவர்களின் தலையில் உரோமம் நரைத்துப் போயும் இருந்ததாம். இதிலிருந்து புகைப்பிடித்தால் உரோமம் கறுப்பாகவும், இல்லையெனில் உரோமம் வெள்ளையாகவும் ஆகிவிடும் என்று ஆய்வாளர் முடிவு செய்யலாம். ஆனால் ஆய்வுக்கு உட்படுத்தப்பட்டவர்களின் வயது பற்றிய புள்ளி விவரத்தைச் சேகரித்துப் பார்த்தபோது உண்மை தெரிய வந்தது. என்னவெனில், புகைப்பிடிப்பவர்களின் சராசரி வயது குறைவாகவும், புகைப்பிடிக்காதவர்களின் சராசரி வயது அதிகமாகவும் இருந்தது. அதிலிருந்து, புகைப்பிடிப்பவர்கள் முடி நரைக்கும் வரை உயிருடன் இருப்பதில்லை என்றும், புகைப்பிடிக்காதவர்கள் முடி நரைக்கும் வரை உயிருடன் இருந்துள்ளார்கள் என்றும் தெரிய வந்ததாம்.

இதேபோல், இன்னுமொரு உதாரணமும் கூறப்படுகிறது. பாலின் நுகர்வின் அளவையும், எலும்பு உருக்கி (Tuberculosis) நோயால் பாதிக்கப்பட்டவர்களின் எண்ணிக்கையையும் பற்றிய புள்ளி விபரங்கள் பல நாடுகளில் சேகரித்தபோது, இவ்விரண்டு மாறிகளுக்கும் இடையே நேரிடை ஒட்டுறவு (Positive correlation) இருந்ததாம். அப்படியானால், பால் அதிகம் நுகரும் நாடுகளில் அந்த நோயால் பாதிக்கப்பட்டவர்களின் எண்ணிக்கை அதிகமாகவும் மற்ற நாடுகளில் அந்த நோய்வாய்ப்பட்டவர்களின் எண்ணிக்கை குறைவாகவும் இருந்ததாம். அப்படியானால், பால்குடிக்காமல் இருப்பதே அந்த நோய் வராமல் இருப்பதற்கு எளிய வழி என்று முடிவு செய்தார்களாம். ஆனால் இந்த எடுத்துக்காட்டிலும், ஆய்வுக்குட்படுத்தப்பட்டவர்களின் வயதினைப் பற்றிய விபரங்கள் சேகரித்தபோது, பால் அதிகம் நுகர்ந்த நாட்டில் மக்கள் அதிகமாக உயிருடன் இருந்ததால், அந்த எலும்பு உருக்கி நோய்க்கு ஆளாகியுள்ளனர் என்றும்,

பால் குறைவாக அருந்திய நாடுகளில் சீக்கிரமே மக்கள் இறந்து விடுவதால் அந்த நோய்க்கு ஆளாகவில்லை எனவும் தெரியவந்ததாம்.

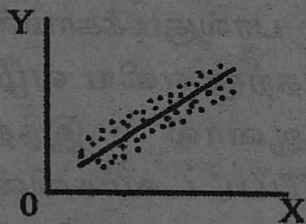
மேற்கூறிய இரண்டு உதாரணங்களும், ஒட்டுறவைக் கொண்டு ஆய்வு செய்யும்பொழுது, முழு விபரங்களும் தெரியாமல் ஆய்வு செய்தால் தவறான முடிவுகளைக் கூற நேரிடும் என்பதை விளக்குகின்றன.

இரண்டு மாறிகளுக்கிடையேயுள்ள உறவு நேரிடையாகவோ (positive) அல்லது எதிரிடையாகவோ (negative) அமையலாம். அந்த உறவு பூரணமாகவோ (perfect) பூரணமில்லாமலோ (imperfect) இருக்கலாம். சிலசமயம் உறவு இருப்பதுபோல் தோன்றி இல்லாமல் இருக்கலாம். இந்தச் சூழ்நிலைக்கு ஒட்டுறவு கண்டுபிடிக்க முயற்சித்தால் அது தவறான போலி (Spurious or nonsense correlation) ஒட்டுறவு ஆகிவிடும். சில சமயங்களில் இரண்டு மாறிகளுக்கிடையே உண்மையாகவே (நேரடியாகவோ மறைமுகமாகவோ) உறவு இருக்கலாம். ஆனாலும் ஆய்வாளர் அவரின் கவனக் குறைவால் அவ்விரு மாறிகளுக்கும் இடையேயுள்ள உறவினைக் கவனிக்காமல் விட்டு விடலாம்.

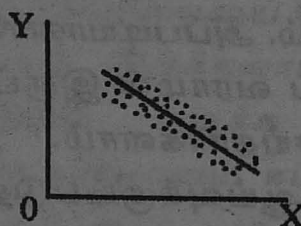
மேலே கூறப்பட்ட சூழ்நிலைகளில் இருந்து, இரண்டு மாறிகளுக்கு இடையேயுள்ள உறவின் தன்மையைச் சரியாக

வரைபடம் - 19

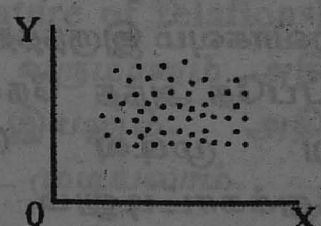
ஒட்டுறவின் வகைகள்



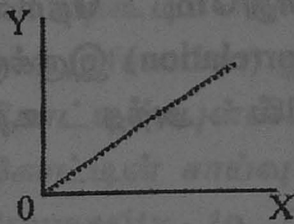
(அ) நேரிடை நேர்கோட்டு ஒட்டுறவு



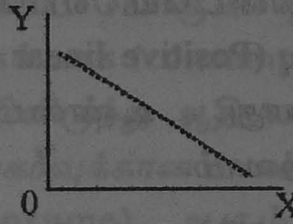
(ஆ) எதிரிடை நேர்கோட்டு ஒட்டுறவு



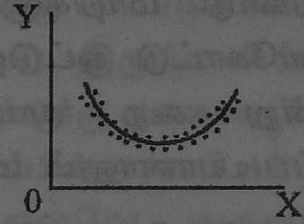
(இ) ஒட்டுறவின்மை



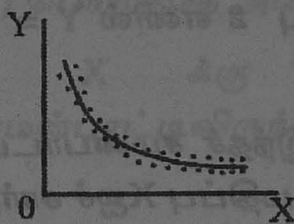
(ஈ) பூரண நேரிடை ஒட்டுறவு



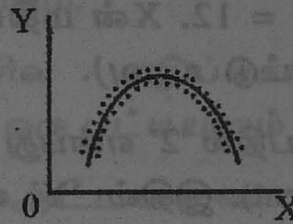
(உ) பூரண எதிரிடை ஒட்டுறவு



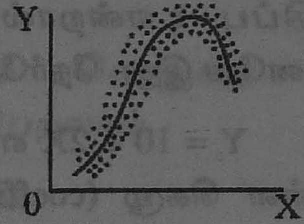
(ஊ) நேர்கோடல்லாத ஒட்டுறவு



(எ) நேர்கோடல்லாத ஒட்டுறவு



(ஏ)நேர்கோடல்லாத ஒட்டுறவு



(ஐ) நேர்கோடல்லாத ஒட்டுறவு

அறிவது சிரமமான பணி என்பதும், இருப்பினும் சரியாக அறிவது ஆய்வாளரின் கடமை என்பதும், அப்படிச் செய்ய வில்லையேல் அந்த ஆய்வின் பலன் குறைவு என்பதும் புலப்படும்.

மேலே வரைந்து கூறப்பட்டுள்ள உறவுகள் எல்லா வகையான ஆய்வுகளிலும் வரலாம்.

நேரிடை நேர்கோட்டு எளிய ஒட்டுறவு (Positive linear simple correlation)

இரண்டு மாறிகளையும் X என்றும் Y என்றும் கொள்வோம். Xன் மதிப்பு பூஜ்யமாக (0) இருக்கும்போது Yன் மதிப்பு 10 என்றும், பிறகு Xன் மதிப்பு ஒவ்வொன்றாகக் கூடும்போது Yன் மதிப்பு இரண்டு இரண்டாகக் கூடுகின்றதென்றும் கொண்டால்,

$Y = 10 + 2X$ என்று சொல்லலாம்.

இரண்டு மாறிகளுக்கிடையே எப்போதுமே நேரிடை நேர்கோட்டு ஒட்டுறவு (Positive linear correlation) இருக்கும் என்று கூற முடியாது. உண்மையில் அந்த உறவு எப்படியானாலும் மாறலாம்.

இப்போது மேலே கொடுக்கப்பட்ட சமன்பாட்டில் Xன் மதிப்பை ஒவ்வொன்றாகக் கூட்டினால் Yன் மதிப்பு இரண்டு இரண்டாகக் கூடுவதைக் காணலாம். உதாரணத்திற்கு Xன் மதிப்பு 1 என்றால் Y = 12. Xன் மதிப்பு 2 எனில் Y = 14. (எனவே இது நேர்கோட்டு உறவு).

Y = 10 + 2X என்பதில் '2' என்பது இந்தச் சமன்பாட்டில் உள்ள கெழு (coefficient). இதில் 'Y' ன் மதிப்பு Xஐச் சார்ந்து உள்ளதால் இச்சமன்பாட்டை Y on X என்று சொல்லலாம். இந்தச் சமன்பாட்டில் Xன் மதிப்பு பூஜ்யமாக இருக்கும்போது Yன் மதிப்பு 10 ஆக இருப்பதால் இதை 'a' என்றும் Yன் மதிப்பை நிர்ணயிக்கும் '2'ஐ b என்றும் சொல்லலாம். இச்சமன்பாடு Y on X ஆக இருப்பதால், இந்த bஐ b_{YX} என்று சொல்லலாம். எனவே, $b_{YX} = 2$; a = 10. இந்த 'b'க்கு வேறு பெயர்களும் உள்ளன. இது சாய்வு (slope) என்றும் முதல் வகை நுண்கெழு (first order differential) என்றும் அழைக்கப்படுகிறது.

இந்த 'a'யும் 'b'யும் இந்தச் சமன்பாட்டின் முக்கிய தன்மைகளைக் குறிப்பதால் இவை பண்பலகுகள் (parameters) என்றும் அழைக்கப்படுகின்றன.

இந்த 'a'யின் மதிப்பு நேரிடையாகவோ (+) எதிரிடையாகவோ (-) இருக்கலாம். Xன் மதிப்பு கூடும்போது அக்கோட்டின் புள்ளிகளிலிருந்து ஆதிக்கு (0க்கு) வரையும் கோட்டின் சாய்வு கூடுவதும் குறைவதும் இந்த 'a'ன் மதிப்பைப் பொறுத்து உள்ளது. 'a'ன் மதிப்பு +ல் இருந்தால், சாய்வு குறைந்து கொண்டே செல்லும்; -ஆக இருந்தால்

சாய்வு கூடிக்கொண்டே போகும். இங்கு கொடுக்கப்படும் சார்பு (function) நுகர்வாக (consumption) இருந்தால் அந்தக் கோட்டின் புள்ளிகளிலிருந்து ஆதி (origin)க்கு வரையப்படும் கோட்டின் சாய்வு நுகர்வுக்கான மனச்சார்பின் சராசரி (Average Propensity to Consume) ஆகும். எனவே, வருமானம் கூடும்போது நுகர்வுக்கான மனச்சார்பின் சராசரி கூடுமா குறையுமா என்பது இங்கு வரும் 'a'ன் மதிப்பை (+ அல்லது -) பொறுத்தே அமையும்.

X க்கு சில மதிப்புக்கள் கொடுத்து ஒரு சமன்பாட்டிலிருந்து ஒரு பட்டியலும் பெறலாம்.

அட்டவணை - 21

Y மற்றும் Xன் மதிப்புக்கள்

X	Y
0	10
1	12
2	14
3	16
4	18
5	20

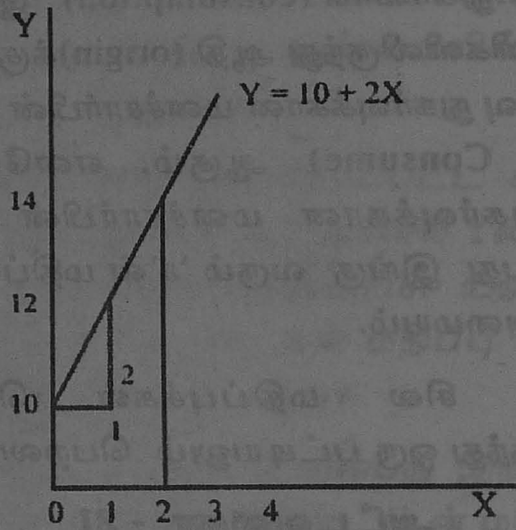
இப்பட்டியலை வரைபடமாக வரைந்தால் வரைபடம் 20 கிடைக்கும்.

கொடுத்துள்ள சமன்பாடான $Y = 10 + 2X$ஐ மாற்றி Xஐ சார்பு மாறியாகவும் (dependent variable) Yஐத் தன்னிச்சை மாறியாகவும் (independent variable) கொண்டால் கீழ்வரும் சமன்பாடு கிடைக்கும்.

$$Y = 10 + 2X$$

$$\frac{1}{2}Y - 5 = X$$

வரைபடம் - 20



வரைபடம் 20இல் 10, Yஅச்சை இடையில் வெட்டுவதால், இதனை (a) Yவெட்டு (Y intercept) எனலாம். அடுத்த சமன்பாட்டில், b (சாய்வு) ½ ஆக உள்ளது. X வெட்டு 5 ஆக உள்ளது. இந்தச் சாய்வுக்கு b_{XY} என்று பெயர் வைக்கலாம். இப்பொழுது கிடைத்துள்ள b_{YX} மற்றும் b_{XY}யிலிருந்து ஒட்டுறவுக் கெழுவினைக் காணலாம்.

$$\text{ஒட்டுறவுக் கெழு} = \sqrt{b_{YX} \cdot b_{XY}}$$

அதாவது b_{YX} மற்றும் b_{XY}ன் பெருக்கல் சராசரியே ஒட்டுறவுக்கெழு ஆகும். இவ்வாறாக Yக்கும் Xக்கும் இடையேயுள்ள ஒட்டுறவு கணக்கிடப்படுகிறது. எனவே இங்குள்ள விபரங்களின்படி

$$r = \sqrt{b_{YX} \cdot b_{XY}} = \sqrt{2 \times 1/2} = \sqrt{1} = 1$$

எனவே, நாம் எடுத்துக்கொண்ட உதாரணத்தில் Yக்கும் Xக்கும் இடையே பூரண நேரிடை நேர்கோட்டு உறவு (perfect, positive, linear relationship) இருக்கிறது.

மேலே உள்ளவாறு சமன்பாடு இல்லாமல், புள்ளி விபரங்கள் மட்டும் கிடைத்தாலும் அவைகளுக்கு இடையேயுள்ள உறவுகளின் அளவையும் (degree of

relationship) உறவுகளின் திசையினையும் (direction of relationship) காணமுடியும்; அதற்கு சில சூத்திரங்களும் உள்ளன. எந்த சூத்திரத்தை எங்கு பயன்படுத்த வேண்டும் என்று கொடுக்கப்பட்ட புள்ளிவிபரங்களின் போக்கினையும், தன்மைகளையும் கருத்தில் கொண்டு முடிவு செய்ய வேண்டும். சரியான புள்ளி விபரங்களுக்குச் சரியான சூத்திரத்தைக் கையாண்டால் விடையினை எளிதாகவும் துல்லியமாகவும் பெறலாம்.

சூத்திரங்கள்

ஒட்டுறவுக் கெழுவினைப் பெற உதவும் சூத்திரங்கள் கீழே தரப்பட்டுள்ளன.

1) $r = \sqrt{b_{YX} \cdot b_{XY}}$

2) $r = \pm \sqrt{\left(\frac{\text{விளக்கப்பட்ட மாறுபாடுகள்}}{\text{மொத்த மாறுபாடுகள்}}\right)}$

$r = \pm \sqrt{\left(\frac{\text{Explained variation}}{\text{Total variation}}\right)}$

$(Y - \bar{Y})$ = Total variation; $Y_c - \bar{Y}$ = Explained variation;

$Y - Y_c$ = unexplained variation)

3) (a) $(X - \bar{X}) = r \frac{s_x}{s_y} (Y - \bar{Y})$ or $x = r \frac{s_x}{s_y} y$ or $x = by$

3) (b) $(Y - \bar{Y}) = r \frac{s_y}{s_x} (X - \bar{X})$ or $y = r \frac{s_y}{s_x} x$

4) $r = \frac{\Sigma xy}{\sqrt{(\Sigma x^2)(\Sigma y^2)}}$ where $x = (X - \bar{X})$; $y = (Y - \bar{Y})$

5) $r = \frac{S_{xy}}{S_x S_y}$

6) $r = \dfrac{N \Sigma XY - \Sigma X \Sigma Y}{\sqrt{[N\Sigma X^2 - (\Sigma X)^2)][N\Sigma Y^2 - (\Sigma Y)^2]}}$

7) $r = \dfrac{N \Sigma d_x d_y - (\Sigma d_x)(\Sigma d_y)}{\sqrt{[N\Sigma d_x^2 - (\Sigma d_x)^2][N\Sigma d_y^2 - (\Sigma d_y)^2]}}$

8) $r = \dfrac{N \Sigma f u_x u_y - (\Sigma f_x u_x)(\Sigma f_y u_y)}{\sqrt{[N\Sigma f_x u_x^2 - (\Sigma f_x u_x)^2)][N\Sigma f_y u_y^2 - (\Sigma f_y u_y)^2]}}$

9) $r = \dfrac{N \Sigma f d_x d_y - (\Sigma f d_x)(\Sigma f d_y)}{\sqrt{[N\Sigma f d_x^2 - (\Sigma f d_x)^2)][N\Sigma f d_y^2 - (\Sigma f d_y)^2]}}$

10) $r = \dfrac{\Sigma (X-\bar{X})(Y-\bar{Y})}{n s_x s_y}$

11) $R = 1 - \dfrac{6 \Sigma d^2}{n^3 - n}$ or $R = 1 - \dfrac{6 \Sigma d^2}{n(n^2 - 1)}$

இதில் R என்பது தர ஒட்டுறவுக்கெழு (rank correlation)

இவற்றில் 'r' ஒட்டுறவுக்கெழு (coefficient of correlation)எனவும்

r^2 தீர்மானக்கெழு (coefficient of determination) எனவும் அழைக்கப்படுகின்றன.

தற்றொடர் ஒட்டுறவு (Autocorrelation)

ஒரு மாறியின் மதிப்பு வேறுவேறு காலங்களிலும் வேறுபடலாம். ஒரு மாறியின் மதிப்பு, உதாரணத்திற்கு, நாட்கள் ஆக ஆகக் கூடிக் கொண்டோ குறைந்து கொண்டோ செல்லலாம். அப்படியிருக்கும்போது அந்த மாறியின் மதிப்பானது காலத்தைப் பொறுத்து அமைகிறது எனலாம். அதை Y = f(T) என்று (அதாவது Yன் மதிப்பு காலத்தைச் சார்ந்துள்ளது) கூறலாம். எடுத்துக்காட்டாக X என்னும்

மாறியின் இன்றுள்ள மதிப்புக்கும், இன்றைக்கு முந்திய காலங்களில் X எடுத்திருந்த மதிப்புக்கும் இடையே உள்ள ஒட்டுறவு, தற்றொடர் ஒட்டுறவு (autocorrelation) எனப்படுகிறது.

அதுபோல, ஒவ்வொரு காலத்திலும் விளக்கப்படாமல் உள்ள பிழைகளை (வேறுபாடுகளை) காலத்தின் அடிப்படையில் அடுக்கி, ஒவ்வொரு பிழையையும் (வேறுபாட்டையும்) அதற்கு முந்தைய காலத்தில் உள்ள பிழையுடன் (வேறுபாட்டுடன்) ஒப்புமைப்படுத்திப் பார்த்து அவைகளுக்கிடையே உறவு இருந்தால் அதுவும் தற்றொடர் ஒட்டுறவு (auto correlation) அல்லது தொடர் ஒட்டுறவு (serial correlation) என அழைக்கப்படுகிறது.

ஒட்டுறவுக்கெழுவின் முக்கியத்துவம்

ஒட்டுறவுக்கெழு -1லிருந்து +1 வரை இருக்கலாம். -1ம் +1ம் முறையே இரண்டு மாறிகளுக்கிடையேயான பூரண எதிரிடை மற்றும் பூரண நேரிடை உடன் தொடர்பினைக் காட்டுகின்றன. உண்மையில் இரண்டு மாறிகளுக்கிடையே பூரண உடன்தொடர்பு இருப்பது என்பது அவ்வளவாக அடிக்கடி நிகழ்வதில்லை. உதாரணத்திற்கு, அட்டவணை 22இல் உள்ள புள்ளி விபரங்களுக்குள் அந்த மாதிரி முழு அல்லது பூரண உடன் தொடர்பு இருக்கலாம். ஆனால் இப்படி ஒரு உறவு முறை இயல்பாகக் காண்பதற்கு வாய்ப்பு குறைவே.

அட்டவணை - 22

Y மற்றும் Xன் மதிப்புக்கள்

X	Y	X	Y
1	1	1	5
2	2	2	4
3	3	3	3
4	4	4	2
5	5	5	1

எனவே அதிகம் காண்பது -1க்கும் +1க்கும் இடையில்தான். இந்த ஒட்டுறவுக்கெழு ஒரு மாறியில் (Y) உள்ள வேறுபாடுகளை எந்த அளவுக்கு இன்னொரு மாறியில் (X) உள்ள வேறுபாடுகள் விளக்குகின்றன என்பதைக் காட்டுகிறது. அல்லது, ஒரு மாறியில் (X) உள்ள வேறுபாடுகள் எந்த அளவுக்கு இன்னொரு மாறியில் ஏற்பட்ட வேறுபாடுகளுக்குப் பொறுப்பாகும் என்பதை ஒட்டுறவுக்கெழு விளக்குகிறது. உதாரணத்திற்கு, Xதன்னிச்சையான மாறி என்றும், Yசார்பு மாறி என்றும், மேலும் ஒட்டுறவுக்கெழுவின் வர்க்கம் 0.8 என்றும் கொண்டால், Yல் ஏற்படுகின்ற 80 விழுக்காடு வேறுபாடுகளை மட்டுமே Xல் ஏற்படுகின்ற வேறுபாடுகள் விளக்குகின்றன என்று பொருள். மீதமுள்ள 20 விழுக்காடு வேறுபாடுகளை வேறு ஏதோ மாறிகள் விளக்குகின்றன. பொதுவாகப் பல முக்கியமான மாறிகளைக் கண்டறிவதிலும், அவற்றைச் சரியாக அளப்பதிலும் சிரமங்கள் இருப்பதால் ஒட்டுறவுக்கெழு -1க்கு அதிகமாகவும் +1க்குக் குறைவாகவும்தான் இருக்கும். பொதுவாக ஒட்டுறவுக்கெழு 0.5ஐவிட அதிகமாக இருந்தால் எடுத்துக்கொண்ட மாறிகளுக்கு இடையேயான உறவைச் சரியாகக் கண்டுபிடித்ததாகப் பொருள் கொள்ளப்படுகிறது. அப்படி 0.5க்கும் அதிகமாக இருந்தால் அது முக்கியத்துவம் (statistiscally significant) பெற்றது என்றும் சொல்கிறார்கள். இந்த **முக்கியத்துவம்** என்ற பதத்தின் புள்ளியியல் பொருள் பின்வரும் வேறு ஓர் அத்தியாயத்தில் விளக்கப்படுகிறது. ஆனாலும் இங்கு அதுபற்றிச் சிறிது சொல்வது புரிதலைக் கொஞ்சம் அதிகப்படுத்தலாம். இரண்டு வேறுவேறு மாறிகளின் முழுமைகளில் (universe, population) இருந்து (X,Y) எடுக்கப்பட்ட மாதிரிகளுக்கிடையேயுள்ள ஒட்டுறவுக் கெழுவினை 'r' என்கிறோம். முழுமைகளுக்கு இடையே உள்ள ஒட்டுறவுக்கெழு ρ (row - ரோ) என்றழைக்கப்படுகிறது; இதனை மாதிரியாக எடுக்கப்பட்ட X மற்றும் Yகளுக்கு

இடையேயுள்ள ஒட்டுறவுக்கெழுவின் (r) மூலம் நிர்ணயிக்கலாம். அப்படிச் செய்யும்போது r முக்கியத்துவம் பெற்றுள்ளது என முடிவு செய்தால், X மற்றும் Y என்ற மாறிகளுக்கிடையே முழுமையிலும் ஒட்டுறவு உள்ளதெனப் பொருள். அப்படியல்லாமல் r முக்கியத்துவம் பெறவில்லை என முடிவு செய்தால் X மற்றும் Yன் முழுமைகளுக்குள்ளும் உறவு இல்லையெனப் பொருள்படும். இதற்கு r, 0.5க்கும் மேல் இருந்தால் ρ முக்கியத்துவம் பெறுகிறது என்றும் r, 0.5க்கும் கீழ் இருந்தால் இரு மாறிகளின் முழுமைகளுக்கு இடையே சரியான உறவில்லை எனவும் முடிவு செய்யலாம் எனச் சிலர் கூறுகின்றனர். இது ஓரளவு சரிதான்; சரியாகச் சொல்லப்போனால் இது ஓரளவுதான் சரி. ρ முக்கியத்துவம் பெறுகிறதா இல்லையா என்பதை r மட்டும் முடிவு செய்வதில்லை. rஐக் கணிக்க எடுத்துக்கொண்ட X, Yன் எண்ணிக்கைகளும் (number of observations) சேர்ந்துதான் முடிவு செய்கின்றன.

உதாரணத்திற்கு மாதிரிகளின் அளவு (sample size) பெரியதாக இருந்தால் (உதாரணத்திற்கு 1000ஆக), மிகச்சிறிய r (உதாரணத்திற்கு 0.2) கூட முக்கியத்துவம் பெற்றுள்ளதாக அமையும். அப்படியில்லாமல், மாதிரிகளின் அளவு சிறியதாக இருந்து (உதாரணத்திற்கு 5), r பெரியதாக (உதாரணத்திற்கு 0.8) இருந்தாலும் r முக்கியத்துவம் பெறாமல் போகலாம்; அதாவது, அந்த மாதிரிகள் (XY) எடுக்கப்பட்ட முழுமைகளுக்கிடையே ஒட்டுறவு (ρ) இல்லை என்றும் முடிவு செய்யப்படலாம். இதை நிர்ணயிப்பதற்கான சூத்திரம்:

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

மேலே கூறப்பட்டுள்ள உதாரணங்களை எடுத்து

$$t = \frac{0.2\sqrt{998}}{\sqrt{1 - 0.04}} = \frac{6.318}{0.98} = 6.45 \qquad (1)$$

$$t = \frac{0.8\sqrt{3}}{\sqrt{1 - 0.64}} = \frac{1.38}{0.6} = 2.3 \qquad (2)$$

முதலாவது *(1)* மதிப்புக்கான அட்டவணை மதிப்பு (t value from the Table) 1.96 (5 விழுக்காடு முக்கியத்தும் அடிப்படையில்). எனவே, இதற்குரிய r முக்கியத்துவம் பெற்றதாக அமைகிறது; இந்த rஐக் கணிப்பதற்கு எடுக்கப்பட்ட மாதிரிகளின் முழுமைகளுக்கிடையே நல்ல உடன்தொடர்பு உள்ளது. இதற்கு மாறாக, இரண்டாவது *(2)* மதிப்புக்கான அட்டவணை மதிப்பு (Table value) *3.18* ஆகும். இதன் மூலம், அந்த மாதிரிகள் எடுக்கப்பட்ட முழுமைகளுக்கு இடையேயான உடன்தொடர்பு முக்கியமானதாக இல்லை எனலாம்.

பட்டியலில் இருக்கும் மதிப்பைவிட, கணக்கிடப்பட்ட t அல்லது z மதிப்பு அதிகமாக இருந்தால் இல்லெனும் எடுகோளை நிராகரிப்போம். இல்லெனும் எடுகோள் (H_0) என்னவெனில் எடுத்துக்கொண்ட மாறிகளின் முழுமைகளுக்கு இடையே உள்ள ஒட்டுறவுக்கெழு (ρ = row) பூஜ்யத்திற்குச் சமம் என்பதாகும். இந்த இல்லெனும் எடுகோளை ஒத்துக் கொண்டோமேயானால் (அதாவது கணிக்கப்பட்ட t, z மதிப்புக்கள் அவற்றின் பட்டியலிடப்பட்ட மதிப்புக்களை விடக் குறைவாக இருந்தால்), எடுத்துக்கொண்ட மாறிகளுக்கிடையே அவற்றின் முழுமைகளில் உறவு இல்லை; அதாவது $\rho = 0$ என்று பொருள். இதில் மாதிரிகளின் அளவு (n) 30க்கு மேல் இருந்தால் 't' பட்டியலும் 'z' பட்டியலும் ஒரே மதிப்புகளைத்தான் தரும். எனவே, பொதுவாக 't'ன் மதிப்பு என்று சமூக அறிவியல்களுக்கான புள்ளியியல் சிப்பங்களில் (Statistical Packages for Social Sciences) காட்டப்பட்டுள்ளது.

மாறாக, மாதிரிகளின் அளவு (n) 30யை விடச் சிறியதாக இருந்தால் 't' பட்டியலில் உள்ள மதிப்புத்தான் பொருத்தமாகும்.

ஒட்டுறவுக்கெழுவின் இரண்டு குணங்கள்

ஒட்டுறவுக்கெழு, ஒரு மாறியில் உள்ள அனைத்து மதிப்புக்களையும் ஒரே எண்ணால் பெருக்கினாலோ வகுத்தாலோ, மாறாது. உதாரணத்திற்கு, ஒரு மாறியின் மதிப்புக்கள் குவிண்டாலில் கொடுக்கப்பட்டிருக்கும்போது அந்த மாறிக்கும் மற்றொரு மாறிக்கும் இடையேயான ஒட்டுறவுக்கெழு என்ன இருந்ததோ, அதேதான், அந்த மாறியின் மதிப்புக்களை குவிண்டாலுக்குப் பதிலாக கிலோ *(100 கிலோ = 1 குவிண்டால்)*வில் மாற்றிக் கண்டுபிடித்தாலும் வரும். இதனை, ஒட்டுறவுக்கெழு அலகினைப் (Unit of measurement) பொறுத்து மாறாது எனலாம். இரண்டாவதாக, ஆதி (Origin)யை மாற்றினாலும், ஒட்டுறவுக்கெழு மாறாது. இணையான மதிப்புக்களை மாற்றாமல் ஓர் இணையை முதலிடத்தில் வைத்து கணித்தால் ஒட்டுறவு கெழு என்னவருமோ, அதேதான் அந்த இணையை மூன்றாவது இடத்தில் வைத்தாலும் நான்காவது இடத்தில் வைத்தாலும் அல்லது வேறு எந்த இடத்தில் வைத்தாலும் வரும். உதாரணத்திற்கு, அட்டவணை 23இல் உள்ள முதல் இரண்டு நிரல்களுக்குமிடையே என்ன ஒட்டுறவுக்கெழு வருகிறதோ அதே அளவு ஒட்டுறவுக்கெழுதான் முதல் மற்றும் மூன்றாம் நிரல்களுக்கிடையேயும் வரும்; அதே அளவு ஒட்டுறவுக்கெழுதான் நான்காம் மற்றும் ஐந்தாம் நிரல்களுக்கிடையேயும் வரும்.

மேலே சொன்னபடி ஒட்டுறவுக்கெழு அமைவதால், மாறிகளின் மதிப்புக்கள் மிகப்பெரிய எண்களாக இருந்து ஒட்டுறவுக்கெழுவினைக் கண்டுபிடிப்பது சிரமமாக இருந்தால், ஒரு மாறியின் அனைத்து மதிப்புக்களையும் ஒரே

எண்ணால் வகுத்துச் சிறியதாக்கிவிட்டு, ஒட்டுறவுக்கெழு கணக்கிடலாம். இன்னொரு மாறியின் மதிப்புக்கள் அதிகமாக இருந்தாலும் மேலே கூறியது போலச் செய்து ஒட்டுறவுக்கெழு காணலாம். மாறிகளின் மதிப்புக்களை மாற்றுவதற்கு முன்புள்ள ஒட்டுறவுக்கெழுவுக்கும் மாற்றிய பின்பு கண்டுபிடிக்கப்பட்ட ஒட்டுறவுக்கெழுவுக்கும் எந்த வித்தியாசமும் இருக்காது.

அட்டவணை - 23

உயரம் (மீட்டரில்) (1)	எடை (கிலோவில்) (2)	எடை (குவிண்டாலில்) (3)	உயரம் (4)	எடை (குவிண்டாலில்) (5)
2	1000	10	6	45
4	2000	20	3	15
7	5000	50	7	50
3	1500	15	4	20
6	4500	45	2	10

மேலும் அட்டவணை 23இல் உள்ள நிரல்களில் முதல் நிரலை X ஆகவும் இரண்டாவது நிரலை Y ஆகக் கொண்டாலும் சரி அல்லது முதல் நிரலை Y ஆகவும் இரண்டாவது நிரலை X ஆகக் கொண்டாலும் சரி ஒட்டுறவுக்கெழு மாறாது. ஆனால் இப்படித் தொடர்புப் போக்கில் (Regression) செய்தால் விடைகள் மாறி வரும். தொடர்புப்போக்குக் கோடுகளின் சாய்வுகள் வித்தியாசமாக வரலாம். ஒட்டுறவுக்கெழு ±1 ஆக இருந்தால் மட்டுமே தொடர்புப்போக்குக் கோடுகளின் சாய்வுகள் 1ஆகவும் சமமாகவும் இருக்கும்.

6.தொடர்புப்போக்கு

(REGRESSION)

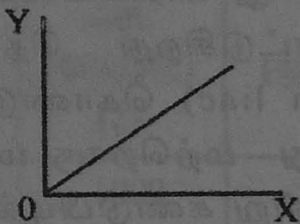
தொடர்புப்போக்கு இரண்டு (அல்லது இரண்டுக்கு மேற்பட்ட) மாறிகளுக்கிடையே உள்ள உறவின் தன்மையையும் (Strength of relationship) உறவின் திசையையும் (direction of relationship) காட்டுவதோடு (ஒட்டுறவுக்கெழு காட்டுவது போல), இரு மாறிகளில் எது காரணம் (Cause) எது விளைவு (effect) என்பதையும் காட்டுகிறது. தொடர்புப் போக்குக் கோட்டைக் (Regression line) கொண்டு, ஒரு மாறியின் மதிப்பு தெரியும்போது மற்றொரு மாறியின் மதிப்பு என்னவாக இருக்கும் என்று கண்டுபிடிக்கலாம். இரண்டு மாறிகள் X,Y என இருந்தால் Yஐச் சார்ந்து X உள்ளதாக வைத்து ஒரு தொடர்புக் கோட்டையும், Xஐச் சார்ந்து Y உள்ளதாக வைத்து ஒரு தொடர்புக் கோட்டையும் கண்டுபிடிக்கலாம். இவற்றில் முதலாவதை X on Y என்றும், இரண்டாவதை Y on X என்றும் கூறலாம். இவ்விரண்டு கோடுகளின் சாய்வுகள் (Slopes) சமமாகவும் வரலாம்; சமமில்லாமலும் வரலாம். அந்தச் சாய்வுகள் சமமில்லாமல் இருந்தால் அந்தக் கோடுகள் ஓரிடத்தில் சந்திக்கும். அப்படிச் சந்திக்கும் இடத்திற்கு நேரே உள்ள Xன் மதிப்பு அதன் கூட்டுச் சராசரியாகவும் அந்த இடத்திற்கு நேரே உள்ள Yன் மதிப்பு Yன் கூட்டுச் சராசரியாகவும் இருக்கும்.

ஒட்டுறவுக்கெழுவுக்கும், தொடர்புப்போக்குக் கோடுகளின் சாய்வுகளுக்கும் நெருங்கிய தொடர்பு உள்ளது. அதை வரைபடங்கள் மூலம் காட்டலாம்.

ஒட்டுறவுக்கெழுவின் வர்க்கம் (coefficient of determination) அல்லது தீர்மானக்கெழு (r^2) கீழ்வரும் விளக்கங்களைத் தருகிறது. *(1)* எந்தளவுக்கு மொத்தப் பிழைகள் (total error, total

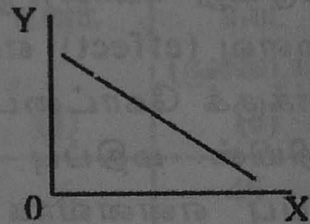
variation) குறைக்கப்படுகின்றன (2)கொடுக்கப்பட்டுள்ள புள்ளிகள் எந்த அளவுக்கு உடன் தொடர்புப் போக்குக் கோட்டுடன் ஒட்டியுள்ளன (3)கொடுக்கப்பட்டுள்ள புள்ளிகள் எந்தளவுக்கு நேர்கோட்டுத் தன்மையைக் (degree of linearity) கொண்டுள்ளன ஆகிய விவரங்களை r^2 தருகிறது.

வரைபடம் - 21
ஒட்டுறவுக் கெழுக்களும் தொடர்புப்போக்குக் கோடுகளின் சாய்வுகளும்



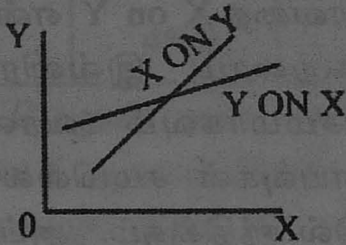
(அ)

r = +1



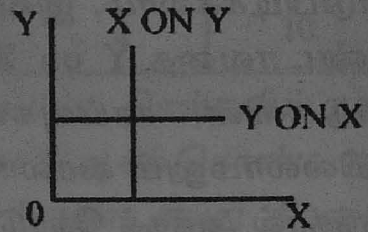
(ஆ)

r = -1



(இ)

-1 < r < +1



(ஈ)

r = 0

பலவகையான கோடுகளுக்குக் கீழ்வரும் சமன்பாடுகளைக் கூறலாம்.

(1) $Y = a + bX \rightarrow$ நேர்கோடு (Straight line)

(2) $Y = a + b_1X_1 + b_2X_1^2 \rightarrow$ பரவளை அல்லது இருபடி வளைகோடு (Parabola or quadratic)

(3) $Y = a + b_1X + b_2X^2 + b_3X^3$ → முப்படி வளைகோடு (Cubic curve)

(4) $Y = a + b_1X + b_2X^2 + b_3X^3 + b_4X^4$ → நாற்படி வளைகோடு (Quartic curve)

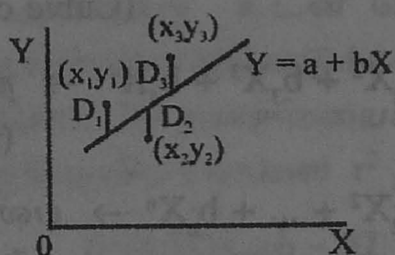
(5) $Y = a + b_1X + b_2X^2 + ... + b_nX^n$ → பலபடி வளைகோடு (n^{th} degree curve)

எந்த வளைகோடு பொறுத்தமாக இருக்கும் என்று கணக்கிடவும் எதைப் பயன்படுத்தினால் சரியான பதில்கள் கிடைக்கும் என்றறியவும் கிடைத்த புள்ளி விபரங்களைச் சிதறல் விளக்கப்படத்தில் பதிந்து பார்க்கலாம்.

சாதாரண மீச்சிறு வர்க்கமுறை
(Method of Ordinary Least Square : OLS)

கிடைக்கப்பட்ட X மற்றும் Y மாறிகளின் மதிப்புக் களுக்கிடையே உள்ள உறவுகளைக் கணக்கிட்டு, கிடைக்காத Yன் மதிப்பையும் (கிடைத்துள்ள Xன் மதிப்பை வைத்து) கிடைக்காத Xன் மதிப்பையும் (கிடைத்துள்ள Yன் மதிப்பின் அடிப்படையில்) கண்டுபிடிக்க சாதாரண மீச்சிறு வர்க்கமுறை பயன்படுகிறது. கணக்கிடப்பட்டிருக்கும் மதிப்புகளுக்கும் உண்மையான மதிப்புகளுக்கும் உள்ள இடைவெளியை D என்று கொண்டு அவற்றின் வர்க்கத்தைக் (D^2) கண்டுபிடித்து அவற்றையெல்லாம் கூட்டினால் (ΣD^2) அது மிகச்சிறியதாக இருத்தல் நலம். ஏனெனில், அந்த இடைவெளிகள் எந்த அளவுக்கு சாதாரண மீச்சிறு வர்க்கமுறையால் கண்டுபிடிக்கப்பட்ட மதிப்புக்களிலிருந்து உண்மையான மதிப்புக்கள் விலகியிருக்கின்றன என்று காட்டுகின்றன. இந்த இடைவெளிகள் (deviation) பிழைகள் (errors)என்றும் மீந்தவை (residuals) என்றும் அழைக்கப்படுகின்றன. இந்தப் பிழைகள் பூஜ்யமாகவோ, நேரிடை எண்களாகவோ (positive) எதிரிடை எண்களாகவோ (negative) இருக்கலாம். இவற்றின் மொத்தம் பூஜ்யமாக *(0)* இருக்க வேண்டும்.

வரைபடம் - 22



மீச்சிறு வர்க்கக்கோடு (Least square line)

மாறிகளான X மற்றும் Yன் மதிப்புக்கள் (X_1,Y_1) (X_2,Y_2) ... (X_n,Y_n) என்பவை கீழே கொடுக்கப்பட்டுள்ள சமன்பாட்டினைக் கொண்டு இருக்கும்.

$Y = a + b_1X_1$ இதில் உள்ள மதிப்புக்கள் தெரியாத 'a' மற்றும் b_1ன் (unknowns) மதிப்புக்களை, கொடுக்கப்பட்டுள்ள சமன்பாட்டிலிருந்து தருவிக்கப்பட்ட கீழ்க்காணும் இரண்டு சமன்பாடுகளை தீர்ப்பதன் (solving simultaneously) மூலம் கண்டுபிடிக்கலாம்.

$$\Sigma Y = na + b_1\Sigma X_1 \; ; \; \Sigma X_1Y = a\Sigma X_1 + b_1\Sigma X_1^2$$

இவ்விரு சமன்பாடுகளும் மீச்சிறு வர்க்கக்கோட்டின் பொதுவான சமன்பாடுகள் என்றழைக்கப்படுகின்றன.

மேலே கொண்டுவந்துள்ளது போல, இரண்டிற்கும் மேற்பட்ட மாறிகள் இருக்கும்போது உள்ள கோடுகளுக்கான சமன்பாடுகளையும், வளைகோட்டு உறவு (non-linear relationship) கொண்டுள்ள மாறிகளுக்கான சமன்பாடுகளையும் தருவிக்க முடியும். உதாரணத்திற்கு, இரு மாறிகள் வளைகோட்டுறவு கொண்டுள்ளபோது

$Y = a + b_1X_1 + b_2X_1^2$க்கான சமன்பாடுகளை

$$\Sigma Y = na + b_1\Sigma X_1 + b_2\Sigma X_1^2 \; ; \; \Sigma X_1Y = a\Sigma X_1 + b_1\Sigma X_1^2 + b_2\Sigma X_1^3$$

$$\Sigma X_1^2Y = a\Sigma X_1^2 + b_1\Sigma X_1^3 + b_2\Sigma X_1^4$$

என்று பெறலாம். மூன்று மாறிகள் உள்ளபோது

$Z = a + b_1X + b_2Y$ என்பதிலிருந்து

$$\Sigma Z = na + b_1\Sigma X + b_2\Sigma Y$$

$$\Sigma XZ = a\Sigma X + b_1\Sigma X^2 + b_2\Sigma XY$$

$$\Sigma YZ = a\Sigma Y + b_1\Sigma XY + b_2\Sigma Y^2$$

என்ற சமன்பாடுகளைப் பெறலாம்.

இதுவரை கூறப்பட்டுள்ள கருத்துக்கள் காலம்சார் தொடர்வரிசையைப் பற்றியும் புரிந்து கொள்வதற்கு மிகவும் உதவும். காலம்சார் தொடர்வரிசையில், X காலத்தைக் குறிக்கும் தன்னிச்சையான மாறியாகவும், Y சார்பு மாறியாகவும் (dependent variable) இருக்கும். புள்ளி விபரங்கள் காலத்தை அடிப்படையாகக் கொண்டு அடுக்கப்பட்டிருக்கும். இவற்றைக் கொண்டு வரையப்படும் கோடுகள் போக்குக் கோடுகள் (trend lines or trend curves) என அழைக்கப்படுகின்றன. இவை எதிர்காலத்தில் நிகழ இருக்கின்ற மாதிரிகளின் மதிப்புக்களைக் கணிக்கப் பயன்படுகின்றன.

புள்ளியியல் சோதனைகள்

இருமாறி எளிய நேர்கோட்டு உறவு காணப் பயன்படும் இரண்டு பண்பலகுகளான aயும் bயும் உண்மையிலேயே புள்ளியியல் முக்கியத்துவம் வாய்ந்தவைதானா என்றறிய, அவற்றைப் புள்ளியியல் சோதனைகளுக்குள்ளாக்குகிறோம். அதற்கு அவற்றின் சராசரிகளும் (Mean) வேறுபாடுகளும் (variances) தேவைப்படுகின்றன. அவற்றைக் கீழ்வருமாறு கணக்கிடலாம்.

$E(\hat{a}) = a$; variance of $\hat{a} = \sigma_u^2 \dfrac{\Sigma X^2}{n\Sigma x^2}$

$E(\hat{b}_1) = b_1$; variance of $\hat{b}_1 = \sigma_u^2 \dfrac{1}{\Sigma x^2}$

இயைபிலா மாறியான என் வேறுபாடு (variance)

$$\sigma_u^2 = \frac{\Sigma e^2}{n-2}$$ ஆகும்.

இவற்றைக் கொண்டு $t_c = \frac{\hat{b}}{SE(\hat{b})}$ என்று கணித்து இந்த t_cயின் மதிப்பு 't'ன் பட்டியல் மதிப்பைவிடக் குறைவாக இருந்தால் $\hat{b}$ பூஜ்யத்திற்குச் சமம் என்றும் புள்ளியியல் முக்கியத்துவம் பெறவில்லை என்றும் கூறலாம். மாறாக, $t_c > t$யாக இருந்தால், $\hat{b}$ புள்ளியியல் முக்கியத்துவம் பெறுகிறது எனலாம். இதனை விரிவாகப் பின்னர் காணலாம்.

பலமாறி ஆய்வு

பொதுவாக அனைத்து மாறிகளுமே (variables) ஒன்றுக்கு மேற்பட்ட மாறிகளின் மதிப்புக்களால் பாதிக்கப்படுகின்றன. அதுபோல் அனைத்து மாறிகளுமே மற்ற சில மாறிகளின் மதிப்புக்களைப் பாதிக்கின்றன. இவ்வாறாக, ஒன்றோடொன்று பின்னிப் பிணைந்து இணைக்கப்பட்ட பல தொடர்புகள் அன்றாடம் காணப்படுகின்றன. ஆனாலும், புள்ளியியலில் பல மாறிகளை ஒரே சமயம் ஆய்வு செய்வதிலும், அவைகளுக்கிடையேயுள்ள உறவுகளை அடையாளம் செய்வதிலும் மிகுந்த சிரமம் இருக்கிறது. மாறிகளின் எண்ணிக்கை கூடும்போது அவற்றின் பண்பலகுகளின் (parameters) எண்ணிக்கையும் கூடும். அவ்வாறுள்ள பண்பலகுகளின் மதிப்புக்களைக் கணிப்பதில் சிரமங்கள் உள்ளன. இப்பொழுது கணினியைப் பயன்படுத்தி கொஞ்சம் எளிதாகக் கணக்குகள் முடிக்கப்படுகின்றன. ஆனாலும், தொடர்புடைய மாறிகளின் எண்ணிக்கை கூடக் கூட கணக்கிடுதலில் உள்ள சிரமங்களும் கூடிக் கொண்டே போகும்.

பல மாறிகளின் மதிப்புக்களைச் சரியாகவும் துல்லியமாகவும் அளந்த பின்னர், அவற்றை சார்பில் (Function) சேர்த்தால் அந்த சார்பின் நம்பகத்தன்மையும் (reliability) துல்லியமும் (Perfection) கூடி அதனுடைய பயன்பாடும் கூடும். கிடைக்கின்றதே என்பதற்காகத் தொடர்பில்லா மாறிகளைச் சார்பில் சேர்த்தாலும் பிரச்சனையைத் தரலாம். மாறுபாடுகள் (variance) கூடலாம். தொடர்புள்ள மாறிகள் கிடைக்கவில்லை எனும் காரணத்திற்காக அவற்றை விட்டுவிட்டு ஆய்வு செய்வதும் முழுமையைத் தராது. எனவே, சரியான மாறிகளைச் சார்பில் (function) சேர்ப்பது அவசியம். அந்த மாறிகளைச் சார்பில் ஒவ்வொன்றாகச் சேர்த்தால், அவ்வாறு சேர்ப்பதின் பயனை நன்றாக அளவிட முடியும். இதனை படிப்படியான தொடர்புப்போக்கு (stepwise regression) எனலாம். சேர்க்கப்படும் ஒவ்வொரு மாறியும் எந்த அளவுக்கு ஒட்டுறவுக் கெழுவின் அளவைக் கூட்டுகிறது என்றும் எந்தளவுக்கு உடன் தொடர்புப் போக்கு பண்பலகுகள் (regression coefficients) பாதிக்கப்படுகின்றன என்றும் பார்க்க முடியும். இவ்வாறாகச் சேர்த்து பலமாறி உடன் தொடர்புப் போக்கினையும் (Multiple regression) பலமாறி ஒட்டுறவுக்கெழு (R, Multiple Correlation) வினையும் காணலாம். மாறிகளின் எண்ணிக்கையைக் கூட்டும்போது மாறிகளின் மதிப்புக்களின் (observations) எண்ணிக்கையையும் கூட்ட வேண்டும். எந்த நேரத்திலும், மாறிகளின் எண்ணிக்கையை (number of variables) விட மாறிகளின் மதிப்புக்கள் (number of observations) எண்ணிக்கையில் அதிகமாக இருக்க வேண்டும். அப்படி இல்லையெனில், அந்த மாறிகளுக்கிடையேயான ஒட்டுறவு மற்றும் உடன் தொடர்புப்போக்குக் கெழுக்களைக் கணக்கிடுவதில் சிரமம் ஏற்படும். அதிகமான மாறிகளைச் சார்பில் சேர்க்கும்போது அதிகமான பண்பலகுகள் வருமாதலால் அவற்றைக் கணிக்க கணினியின் துணை அவசியமாகிறது. இன்று சிரமமான புள்ளி விபரங்களை

ஆய்வுக்குட்படுத்தி முடிவுகளைக் காண பல பொதிமங்கள் (packages) வந்துள்ளன. அவற்றில் எக்ஸெல் (excel), எஸ்.பி.எஸ்.எஸ். (SPSS - Statistical Package for Social Sciences) ஆகியவை குறிப்பிடத்தக்கவை.

மாறிகளை ஒவ்வொன்றாக சார்பில் (function) சேர்த்துக் கொண்டே செல்லும்போது ஒட்டுறவுக்கெழு கூடுவதற்கு வாய்ப்பு உள்ளது. இது கணிதவியல் குணாதிசயம். ஆனால் சேர்க்கப்படும் மாறி உண்மையிலேயே எந்த அளவுக்கு சார்பு மாறியில் (dependent variable) காணும் வேறுபாடுகளை (variation) விளக்குகிறது (explains) அல்லது அந்த வேறுபாடுகளுக்குக் காரணமாக இருக்கிறது என்று காண ஒட்டுறவுக்கெழு பொருந்தாது. ஏனெனில், இந்த ஒட்டுறவுக்கெழு மாறிகளின் எண்ணிக்கையைக் கூட்டினாலேயே கூடும் தன்மை கொண்டது. இந்த பிரச்சனையைச் சரி செய்ய, சரிசெய்யப்பட்ட தீர்மானக்கெழு (adjusted coefficient of determination = $\overline{R}^2$) கணக்கிடப்படுகிறது. அதற்கான சூத்திரம்

$$\overline{R}^2 = 1 - \left[(1-R^2)\,\frac{n-1}{n-k}\right] \text{ அல்லது } 1 - \left[\frac{\Sigma e^2/(n-k}{\Sigma y^2/(n-1)}\right]$$

இது நேரிடையாகவோ (+), எதிரிடையாகவோ (-) வரலாம். ஆனால் R^2 எதிரிடையாக வராது. R^2 பல மாறிகளுக்கிடையே உள்ள ஒட்டுறவுக் கெழுவினைக் குறிக்கிறது. n மாறிகளின் மதிப்புக்களின் எண்ணிக்கையையும் k மாறிகளின் எண்ணிக்கையையும் குறிக்கின்றன. மாதிரியின் அளவு (n) மிகப்பெரியதாக இருந்தால் R^2 ம் $\overline{R}^2$ம் சமமாக இருக்கலாம். $\overline{R}^2$ன் மதிப்பு எதிரிடையாக (-) இருந்தால் அது பூஜ்யம் (0) என்று பொருள்.

ஒரு சார்பில் எத்தனை மாறிகள் வேண்டுமானாலும் இருக்கலாம். இரண்டு விளக்கும் (predictor, explanatory or

independent) மாறிகளும் ஒரு விளக்கப்படும் (outcome, explained or dependent) மாறியும் இருந்து அவற்றிற்கிடையே நேர்கோட்டு உறவு இருந்தால் அந்தச் சார்பு கீழ்வருமாறு இருக்கும்.

$$Y = a + b_1X_1 + b_2X_2 + u$$

இதில் u என்பது இயைபிலா மாறி (Random variable). இது பிழை (error) எனவும், தொந்தரவு (disturbance term) எனவும் அழைக்கப்படுகிறது.

b_1, b_2 என எழுதினால் அவை உண்மையான பண்பலகுகள் எனப் பொருள். அவ்வாறன்றி $\hat{b}_1$, $\hat{b}_2$ என இருந்தால் உண்மையான பண்பலகுகளின் மதிப்பீடு (estimate or estimator) எனப் பொருள். அப்படி எழுதும்போது

$\hat{Y} = \hat{a} + \hat{b}_1X_1 + \hat{b}_2X_2$ எனவும் எழுதலாம்.

$$\hat{a} = \hat{Y} - \hat{b}_1\overline{X}_1 - \hat{b}_2\overline{X}_2$$

$$\hat{b}_1 = \frac{(\Sigma x_1y)(\Sigma x_2^2) - (\Sigma x_2y)(\Sigma x_1x_2)}{(\Sigma x_1^2)(\Sigma x_2^2) - (\Sigma x_1x_2)^2}$$

$$\hat{b}_2 = \frac{(\Sigma x_2y)(\Sigma x_1^2) - (\Sigma x_1y)(\Sigma x_1x_2)}{(\Sigma x_1^2)(\Sigma x_2^2) - (\Sigma x_1x_2)^2}$$

$$R^2_{Y.X_1X_2} = \frac{\Sigma(\hat{Y} - \overline{Y})^2}{\Sigma(Y - \overline{Y})^2} = 1 - \frac{\Sigma e^2}{\Sigma y^2}$$

அல்லது

$$R^2_{Y.X_1X_2} = \frac{\hat{b}_1\Sigma yx_1 + \hat{b}_2\Sigma yx_2}{\Sigma y^2}$$

அல்லது

$$R^2 = 1 - \left[\frac{\Sigma e^2}{\Sigma y^2}\right]$$

பண்பலகுகளின் மதிப்பீடுகளின் சராசரியும் (Mean) மாறுபாடும் (variance)

பண்பலகுகளின் (parameters) மதிப்பீடுகளின் (estimates) சராசரியும் மாறுபாடும் அவற்றைச் சோதித்துப் பார்க்க (Test of significance) உதவும். அவற்றைக் கீழ்வருமாறு தரலாம்.

$$E(\hat{a}) = a \qquad E(\hat{b}_1) = b_1 \qquad E(\hat{b}_2) = b_2$$

$$\text{Variance of } a = \sigma_u^2\left[\frac{1}{n} + \frac{\bar{X}_1^2\Sigma x_2^2 + \bar{X}_2^2\Sigma x_1^2 - 2\bar{X}_1\bar{X}_2\ \Sigma x_1x_2}{\Sigma x_1^2\,\Sigma x_2^2 - (\Sigma x_1x_2)^2}\right]$$

$$\text{Varianceof } b_1 = \sigma_u^2\ \frac{\Sigma x_2^2}{\Sigma x_1^2\,\Sigma x_2^2 - (\Sigma x_1x_2)^2}$$

$$\text{Varianceof } b_2 = \sigma_u^2\ \frac{\Sigma x_1^2}{\Sigma x_1^2\,\Sigma x_2^2 - (\Sigma x_1x_2)^2}$$

இங்கு $\sigma_u^2 = \frac{\Sigma e^2}{n-k}$, k என்பது பண்பலகுகளின் அல்லது மாறிகளின் எண்ணிக்கை. உதாரணத்திற்கு, இரண்டு தன்னிச்சையான (independent) மாறிகளும் ஒரு சார்பு மாறியும் (dependent variable) இருந்தால், அங்கு மூன்று பண்பலகுகள் Parameters, a, b_1, b_2) இருப்பது என்பது உறுதி. எனவே, அங்கு k = 3.

மூன்று மாறிகளுக்கிடையேயான உறவு

அட்டவணை - 24

தேவை, விலை, செலவு ஆகியவற்றிற்கிடையேயான உறவுகளை அளவிடுதல்

தேவையின் அளவு எண்ணிக்கையில் (Y)	விலை ரூபாயில் (X_1)	வருமானம் ரூபாயில் (X_2)
100	5	1000
75	7	600
80	6	1200
70	6	500
50	8	300
65	7	400
90	5	1300
100	4	1100
110	3	1300
60	9	300

ஒரு பொருளின் தேவையின் அளவு (Y) அந்தப் பொருளின் விலையையும் (X_1) நுகர்வோரின் வருமானத்தையும் (X_2) பொறுத்துள்ளது என்று மேலே உள்ள அட்டவணை 24இல் கொடுக்கப்பட்டுள்ளது. இதில்,

$n = 10$ $\quad\quad \bar{X}_2 = 800$

$\Sigma Y = 800$ $\quad\quad \Sigma(Y - \bar{Y}) = \Sigma y = 0$

$\Sigma X_1 = 60$ $\quad\quad \Sigma(X_1 - \bar{X}_1) = \Sigma x_1 = 0$

$\Sigma X_2 = 8000$ $\quad\quad \Sigma(X_2 - \bar{X}_2) = \Sigma x_2 = 0$

$\bar{Y} = 80$

$\bar{X}_1 = 6$

$\Sigma(Y - \bar{Y})^2 = \Sigma y^2 = 3450$

$\Sigma(X_1 - \bar{X}_1)^2 = \Sigma x_1^2 = 30$

$\Sigma(X_2 - \bar{X}_2)^2 = \Sigma x_2^2 = 15,80,000$

$\Sigma yx_1 = - 300$

$\Sigma yx_2 = 65,000$

$\Sigma x_1 x_2 = -5,900$

இந்த விபரங்கள் அட்டவணை 24இல் இருந்து பெறப்பட்டவை. இவற்றிலிருந்து கீழ்வரும் மதிப்புக்களைப் பெறலாம்.

$$\hat{b}_1 = \frac{(-300)(1580000) - (65000)(-5900)}{(30)(1580000) - (5900)^2} = -7.1882$$

$$\hat{b}_2 = \frac{(65000)(30) - (-300)(-5900)}{(30)(1580000) - (5900)^2} = 0.0143$$

எனவே,

$$\hat{b}_0 = \bar{Y} - \hat{b}_1 \bar{X}_1 - \hat{b}_2 \bar{X}_2$$

$$= 80 - (-7.19)(6) - (0.0143)(800) = 111.69$$

(இது $\hat{a}$ என்றும் சில புத்தகங்களில் $\hat{b}_0$ என்றும் குறிக்கப்படுகிறது)

இதிலிருந்து $Y = 111.69 - 7.1882X_1 + 0.0143X_2$ என்றும் என்னென்ன விலையில் (X_1) எந்தளவு வருமானத்தில் (X_2) எவ்வளவு தேவையிருக்கும் (Y) என்றும் கணக்கிட்டுச் சொல்லி விடலாம்.

இதற்குத் தீர்மானக்கெழு (coefficient of determination) வினையும் (R^2) கணக்கிடலாம்.

$$R^2 = \frac{\hat{b}_1\Sigma yx_1 + \hat{b}_2\Sigma yx_2}{\Sigma y^2} = 0.894$$

இதிலிருந்து தேவையில் ஏற்படும் வேறுபாடுகளில் 89.4 விழுக்காடு வேறுபாடுகளை விலையிலும் வருமானத்திலும் ஏற்படும் வேறுபாடுகள் விளக்குகின்றன என்றும் மீதமுள்ள 10.6 விழுக்காடு வேறுபாடுகளை வேறு ஏதோ காரணிகள் (மாறிகள்) விளக்குகின்றன என்றும் பொருள் கொள்ளலாம்.

$\hat{b}_1$, $\hat{b}_2$ ஆகியவைகளின் திட்டப்பிழைகளை (standard error) மதிப்பிடுவதற்கு σ^2_u மதிப்பீடு தேவைப்படுகிறது.

$$\hat{\sigma}^2_u = \Sigma e^2 / n\text{-}k$$

$$R^2 = 1 - \left[\frac{\Sigma e^2}{\Sigma y^2}\right]$$

இதிலிருந்து $\Sigma e^2 = \Sigma y^2 \ (1\text{-}R^2)$

$$\Sigma e^2 = (3,450)\ (0.106) = 365.7$$

$$\text{எனவே, } \hat{\sigma}^2 u = \frac{365.7}{10\text{-}3} = 52.27$$

$(\hat{b}_1)$ ன் மாறுபாடு (variance) $= 52.24 \dfrac{1580000}{12590000} \approx 6.53$

$(\hat{b}_2)$ ன் மாறுபாடு (variance) $= 52.24 \dfrac{30}{12590000} = 0.0001$

$(\hat{a})$ ன் மாறுபாடு (variance) $= 553.69$

இந்த மதிப்பீடுகளின் திட்டப்பிழைகள்

$$s(\hat{b}_1) = 2.55; \quad s(\hat{b}_2) = 0.01; \quad s(\hat{a}) = 23.5$$

கொடுக்கப்பட்டுள்ள மூன்று மாறி (Three variable) தொடர்புப் போக்கினைக் கீழ்வருமாறு எழுதலாம்.

$$\hat{Y} = 111.7 - 7.19\, X_1 + 0.014X_2$$

$$s(b)\ (23.5) \quad (2.55) \quad (0.01) \qquad R^2 = 0.894$$

$$t^*\ (4.75) \quad (-2.8) \quad (1.28) \qquad (2.365)$$

கிடைக்கப்பெற்றுள்ள முடிவுகளிலிருந்து 5 சதவீத அளவில் $\hat{a}$ யும் $\hat{b}_1$ யும் புள்ளியியல் முக்கியத்துவம் பெற்றுள்ளன. ஆனால் $\hat{b}_2$ புள்ளியியல் முக்கியத்துவம் பெறவில்லை.

பகுதி ஒட்டுறவுக்கெழு (Partial Correlation coefficient)

கொடுக்கப்பட்டுள்ள மாறிகளில், மற்ற மாறிகள் மாறவில்லை என வைத்துக்கொண்டு ஏதேனும் இரண்டு மாறிகளுக்கிடையே உள்ள உறவினை மட்டும் சொல்வது பகுதி ஒட்டுறவுக்கெழுவாகும். முதலில் எடுத்துக் கொண்டுள்ள உதாரணத்திற்கிணங்க, ஒரு பொருளின் விலைக்கும் அதனுடைய தேவைக்கும் எதிரிடை உறவு இருக்கலாம். ஆனால் இந்த இரண்டு மாறிகளும் (விலை, தேவை) வருமானத்தின் அளவாலும் பாதிக்கப்படுகின்றன. மக்களின் வருமானம் கூடும்போது தேவையும் கூடலாம்; விலையும் கூடலாம். எனவே விலைக்கும் தேவைக்கும் உள்ள உறவை மட்டும் பார்த்துக் கொள்கையளவில் முடிவெடுப்பது தவறான முடிவுகளைத் தந்துவிடலாம். விலைக்கும் தேவைக்கும் உள்ள உறவினை வருமானம் மாற்றி விடலாம். எனவே, விலைக்கும் தேவைக்கும் உள்ள உறவினை அளவிடும்போது, வருமானத்தில் ஏற்படுகின்ற மாற்றங்களையும் கவனத்தில்கொள்ள வேண்டும். வருமானம் மாறவில்லை என வைத்துக்கொண்டு விலைக்கும் தேவைக்கும் உள்ள உறவினை மதிப்பிடுவதன் மூலம் வருமானத்தைச் சேர்க்காததால் வருகின்ற குறைகளைத் தவிர்க்கலாம். இவ்வாறாகக் கணக்கிடப்படும் பகுதி ஒட்டுறவுக் கெழு (Partial Correlation Coefficient) பல மாறிகளுக்கிடையேயுள்ள எளிய

ஒட்டுறவுக்கெழுவை (Simple Correlation Coefficient) வைத்தே தீர்மானிக்கப்படுகிறது.

r_{12} என்றால் X_1க்கும் X_2க்கும் இடையேயுள்ள ஒட்டுறவுக்கெழு

r_{13} என்றால் X_1க்கும் X_3க்கும் இடையேயுள்ள ஒட்டுறவுக்கெழு

r_{23} என்றால் X_2க்கும் X_3க்கும் இடையேயுள்ள ஒட்டுறவுக்கெழு

$r_{12.3}$ என்றால் X_3 மாறவில்லை என்று வைத்துக்கொண்டு
X_1க்கும் X_2க்கும் இடையேயுள்ள பகுதி ஒட்டுறவுக்கெழு

$$r_{12.3} = \frac{r_{12} - (r_{13})(r_{23})}{\sqrt{(1-r^2_{13})(1-r^2_{23})}}$$

$$r_{13.2} = \frac{r_{13} - (r_{12})(r_{23})}{\sqrt{(1-r^2_{12})(1-r^2_{23})}}$$

நேர்கோட்டுத் தொடர்புப்போக்கு அநுமானங்கள் (Assumptions of the Linear Stochastic Regression Model)

நேர்கோட்டுத் தொடர்புப்போக்கு சில அநுமானங்களின் அடிப்படையில் கணக்கிடப்படுகிறது. இந்த அநுமானங்கள் மூவகைப்படும். முதல்வகை u என்ற இயைபிலா மாறியின் பரவலை அடிப்படையாகக் கொண்டது.

இரண்டாம் வகை, இயைபிலா மாறியான uக்கும் மற்ற தீர்மானிக்கும் மாறிகளுக்கும் (explanatory variables) இடையேயுள்ள உறவின் தன்மையைப் பொறுத்தது. மூன்றாம் வகை, தீர்மானிக்கும் மாறிகளுக்கிடையேயுள்ள உறவின் முறைகளைப் பொறுத்தது.

முதல்வகை அனுமானங்கள்

முதல்வகை அனுமானங்கள் பண்பலகுகளை மதிப்பீடு செய்வதற்கு மிகவும் முக்கியமானவையாக இருக்கின்றன.

அனுமானம் 1 : 'u' என்பது இயைபிலா உண்மை மாறி (random real variable). uன் மதிப்பு ஒரு காலத்தில் என்ன வேண்டுமானாலும் இருக்கலாம்; நேரிடையாகவோ (+) எதிரிடையாகவோ (-), பூஜ்யமாகவோ (0) இருக்கலாம். ஒவ்வொரு மதிப்பும் ஒரு குறிப்பிட்ட நிகழ்தகவின் அடிப்படையில் அமைகிறது.

அனுமானம் 2 : 'u'ன் சராசரி எந்த ஒரு குறிப்பிட்ட காலத்திலும் பூஜ்யமாக இருக்கும். ஒவ்வொரு Xன் மதிப்புக்கும் 'u' எந்த மதிப்பு வேண்டுமானாலும் எடுக்கலாம் என்பதனால், கொடுக்கப்படும் ஒரு Xன் மதிப்புக்கு, எல்லா 'u'க்களையும் கூட்டினால் அதன் மொத்தம் பூஜ்யமாக இருக்கும்; அதனால், 'u'ன் சராசரியும் பூஜ்யமாக இருக்கும்.

இந்த அனுமானப்படி, $Y = a + b_1X$ என்று கொள்ளலாம். 'u'ஐ இதில் சேர்க்க வேண்டியதில்லை.

அனுமானம் 3 : 'u'ன் மாறுபாடு ஒவ்வொரு காலத்திலும் நிலையாக மாறாமல் இருக்கும் (Variance of u is constant in each period). எல்லா Xன் மதிப்புக்களுக்கும், u தன்னுடைய சராசரியைச் சுற்றி ஒரே அளவிலான பரவலைக் கொண்டிருக்கும் (Homoscedasticity அனுமானம்).

அனுமானம் 4 : 'u' இயல்நிலைப் பரவலை (normal distribution) கொண்டுள்ளது.

இந்த நான்கு அனுமானங்களையும் சுருக்கமாக $u \sim N(0, \sigma^2_u)$ எனச் சொல்லலாம்.

அனுமானம் 5 : 'u' பெறும் பல்வேறு மதிப்புக்கள் (u_iu_j) தன்னிச்சையானவை. அதாவது, u_iu_jன் கூட்டு மாறுபாடு (covariance) பூஜ்யமாக இருக்கும். ஒரு காலத்தில் 'u' பெறும் மதிப்பும் மற்றுமொரு காலத்தில் 'u' பெறும் மதிப்பும் சாராமல்

(independent) இருக்கும். சார்ந்திருந்தால் அதற்குப் பெயர் தொடர் ஒட்டுறவு (serial correlation or autocorrelation) ஆகும்.

இரண்டாம்வகை அநுமானம்

அநுமானம் 6 : இடையூறு உறுப்பாகக் (disturbance term) கருதப்படும் 'u' க்கும் தீர்மானிக்கும் மாறி (explanatory variable)க்கும் எந்தவித உறவும் இல்லை; அவை இரண்டும் ஒரேவிதமாக வேறுபடமாட்டா. இந்த அநுமானத்தைச் சுருக்கமாக,

Cov (Xu) = 0 எனலாம்.

உதாரணத்திற்கு, ஒரு கடையில் ஒரு பொருளின் விலை (எடுத்துக்காட்டாக சலவைக் கட்டியைக் கொள்ளலாம்) சில நாட்களுக்கு மாறாமலே இருக்கிறது. ஆனாலும் அதன் விற்பனையில் ஒவ்வொரு நாளும் மாற்றம் இருக்கலாம். ஒரு நாள் இரண்டும், மற்றொரு நாள் நான்கும் இன்னொரு நாள் ஐந்தும் விற்பனையாகி இருக்கலாம். ஆனால், இந்த நாட்களில் சலவைக் கட்டியின் விலையில் மாற்றமில்லாமல் இருந்திருக்கலாம். எனவே இங்கு சலவைக்கட்டியின் விற்பனையில் ஏற்பட்ட மாற்றங்களுக்கு சலவைக்கட்டியின் விலை காரணமாக இருக்க வாய்ப்பில்லை. வேறு ஏதாவது நிகழ்வுகள், விற்பனை மாற்றங்களுக்குக் காரணமாக இருந்திருக்கலாம்; அவற்றை இயைபிலா பாதிப்புக்கள் (random influences) எனலாம். அப்படியிருந்தால், X க்கும் u க்கும் எந்த உறவும் இருக்காது.

மூன்றாம்வகை அநுமானங்கள்

அநுமானம் 7 : தீர்மானிக்கும் மாறிகளான $X_1, X_2 \ldots, X_n$ ஆகியவை பிழை இன்றி அளக்கப்பட்டுள்ளன.

அநுமானம் 8 : தீர்மானிக்கும் மாறிகளான (explanatory variables) $X_1, X_2 \ldots, X_n$ ஆகியவற்றிற்கிடையே பூரண நேர்கோட்டு உறவு

(perfect linear correlation) இல்லை. அவைகளுக்கிடையே பூரணமற்ற உறவு இருக்கலாம்.

அநுமானம் 9 : அனைத்து மாறிகளும் சரியாக அளவிடப்பட்டு மொத்தமாக்கப்பட்டன. உதாரணத்திற்கு ஒரு நுகர்வுச் சார்பில் (consumption function : $c = a + b_1 Y + u$), நுகர்வும் (c), வருமானமும் (Y) பல வீடுகளினுடைய கூடுதலாக இருக்கலாம். அப்படி அவை கூட்டப்பட்டிருந்தால், அவை சரியான முறையைக் கடைப்பிடித்து கூட்டப்பட்டிருக்க வேண்டும்.

அநுமானம் 10 : தீர்மானிக்கும் (explanatory) மாறிகளுக்கிடையேயுள்ள உறவுமுறைகளும் சரியாக நிர்ணயிக்கப்பட்டுள்ளன. தேவையான அனைத்து தீர்மானிக்கும் மாறிகளும் சார்பில் சேர்க்கப்பட்டுள்ளன.

மேலே கூறப்பட்டுள்ள அநுமானங்கள் தவறினாலோ மீறப்பட்டாலோ (violation of assumptions), எதிர்பார்க்கப்படும் சரியான விளைவுகளைப் பெறுவதில் சிரமங்கள் ஏற்படலாம். என்ன சிரமங்கள் ஏற்படலாம்? அவற்றை எப்படிக் கண்டுபிடிப்பது? அவற்றினால் என்ன விளைவுகள் நேரிடலாம்? அவற்றைத் தவிர்க்க அல்லது சரி செய்ய என்ன செய்ய வேண்டும்? என்பன போன்ற கேள்விகளுக்குப் பதில் தெரிவது அவசியம். அப்படித் தெரிந்தால்தான் சரியான, பலனுள்ள ஆய்வுகளைச் செய்ய முடியும். இதுபற்றி முழுவிபரம் தெரிய எகனாமெட்ரிக்ஸ் (Econometrics) புத்தகங்களைப் பயன்படுத்தலாம். அல்லது டேரோ யமனே (TARO YAMANE) எழுதியுள்ள புள்ளியியல் : ஓர் அறிமுக ஆய்வு (Statistics : An Introductory Analysis), மூன்றாவது பதிப்பு, ஹார்ப்பர் பன்னாட்டு பதிப்பு (Harper International Edition, New York, 1973) புத்தகத்தில் பக்கம் 901 முதல் 1019 வரை பார்க்கலாம்.

மேலே சொல்லப்பட்டுள்ள விவரங்கள் முழுமையையும் விளக்க இடமில்லை என்பதால், அவற்றின் ஒரு சுருக்கம் மட்டும் இங்கு தரப்படுகிறது.

இயைபிலா இடையூறு மாறியான (Stochasitc disturbance variable) eக்களுக்குள் உள்ள ஒட்டுறவு பூஜ்யமாக இருக்க வேண்டும். அப்படியில்லையெனில் அதற்கு தொடர்பு ஒட்டுறவு (serial correlation) அல்லது தற்றொடர் ஒட்டுறவு (auto correlation) பிரச்சனை என்று பெயர். இதைச் சுருக்கமாக $E(e_ie_j) = 0$ என இருக்க வேண்டும் என்றும், அப்படியில்லாது $E(e_ie_j) \neq 0$ என இருந்தால் அங்கு தொடர் ஒட்டுறவுப் பிரச்சனை இருக்கிறது என்றும் கூறலாம்.

இத்தொடர் ஒட்டுறவு எந்த அளவுக்கு தொடர்புப் போக்கின் பண்பலகுகளின் மதிப்பீடுகளைப் (estimations) பாதிக்கும் என்பது பற்றி எழுத இங்கே இடம்போதாது. எனவே, முடிவுகளை மட்டும் மிகச்சுருக்கமாக இங்கே படிக்கலாம்.

முதலாவதாக, தொடர் ஒட்டுறவு இருந்தாலும், அது உடன் தொடர்போக்கு பண்பலகுகளின் அளவையும் மதிப்பீடுகளையும் (estimates or estimators) பாதிக்காது. ஆனால் அந்தப் பண்பலகுகளின் மாறுபாட்டினை (variance) குறைத்துக்காட்டும். எனவே, இது பூஜ்ய எடுகோளை (Null Hypothesis) மறுப்பதற்கான வாய்ப்பினை உருவாக்கும். இப்பிரச்சனையைத் தவிர்க்க, முதலில் புள்ளி விபரங்களுக்கிடையே தற்றொடர் ஒட்டுறவு இருக்கிறதா எனக் கண்டறிய வேண்டும். அதைக் கண்டறிய டர்பின்-வாட்சன் (DURBIN-WATSON Test) சோதனை பயன்படுத்தப்படுகிறது.

டர்பின் - வாட்சன் சோதனை (DURBIN-WATSON Test)

உதாரணத்திற்கு இருபது ஆண்டுகளுக்கு ஒரு நாட்டின் வருமானம் (X_1) அந்த நாட்டில் உள்ள நுகர்வோர் எண்ணிக்கை (X_2) அங்கு விற்பனையான தொலைக்காட்சிப் பெட்டிகளின் எண்ணிக்கை (Y) ஆகியவை பற்றிய புள்ளி விபரங்கள் கிடைத்துள்ளன எனக் கொள்வோம். அதற்கு முழுத்தொகை உடன்தொடர்புக்கோடு (Population regression line) : $Y = a + b_1X_1 + b_2X_2 + e$ என்றும்,

அதன் உடன் தொடர்புக் கோட்டின் மதிப்பீடு (estimated regression line)

$Y = a + b_1X_1 + b_2X_2$ என்றும் இருக்கும். மேலே கொடுத்துள்ளதில், e என்பது இயைபிலா இடையூறுப் பகுதி (stochastic disturbance term, white noise). இதற்கு அனைத்தையும் உள்ளடக்கியது (catch all term) என்றும் பெயர். ஏனெனில், இது, X_1 மற்றும் X_2வைத் தவிர, மற்ற காரணிகள் (factors) Yஐப் பாதித்த அளவினைத் தன்னகத்தே கொண்டுள்ளது. சென்ற ஆண்டு தொலைக்காட்சிப் பெட்டிகளின் விற்பனை இந்த ஆண்டு தொலைக்காட்சிப் பெட்டிகளின் விற்பனையை ஓரளவுக்குப் பாதிக்கலாம். ஆனால் இது ஒரு தனி மாறியாக மேலேகொடுத்த சார்பில் (function) எடுத்துக் கொள்ளப்பட வில்லை. எனவே e_t இந்த ஆண்டின் தொலைக்காட்சிப் பெட்டிகளின் விற்பனையையும், e_{t-1}, சென்ற ஆண்டின் தொலைக்காட்சிப் பெட்டிகளின் விற்பனையையும் உள்ளடக்கியிருக்கும். ஆனால் இவையிரண்டும் தன்னிச்சையாக நடப்பவை அல்ல; அவைகளுக்கிடையே நெருங்கிய தொடர்பு உள்ளது. ஆனாலும், இவைகளுக்கிடையே எந்த உறவும் இல்லை என அனுமானம் செய்யப்பட்டுள்ளது. இது சரியல்ல.

முன்னர் கூறிய பிழைப்பகுதியைக் (error term) கண்டறிய டர்பினும் வாட்சனும் சேர்ந்து ஒரு சூத்திரத்தை உருவாக்கியுள்ளனர். அது,

$$d = \frac{\sum_{t=2}^{n}(e_t - e_{t-1})^2}{\sum e_t^2}$$

இந்த dயின் மதிப்பு பூஜ்யமாகவோ அல்லது மிகச்சிறியதோகவோ இருந்தால், அங்கு நேரிடை தற்றொடர் ஒட்டுறவு (Positive autocorrelation or serial correlation) இருக்கிறதென்று பொருள். டர்பினும் வாட்சனும் இதைத் தீர்மானிக்க ஒரு அட்டவணை கொடுத்துள்ளார்கள். அந்த அட்டவணையில் dயின் கீழ்ப்பகுதி மதிப்பும் (d_L = d lower), dயின் மேல்பகுதி மதிப்பும் (d_U = d upper) 5%, 2.5% மற்றும் 1% முக்கிய அளவில் (level of significance) கொடுக்கப்பட்டுள்ளன. அவற்றின் உதவி கொண்டு தொடர் ஒட்டுறவின் தன்மைகளைக் (நேரிடையா, எதிரிடையா அல்லது பூஜ்யமா) கண்டுபிடிக்கலாம்.

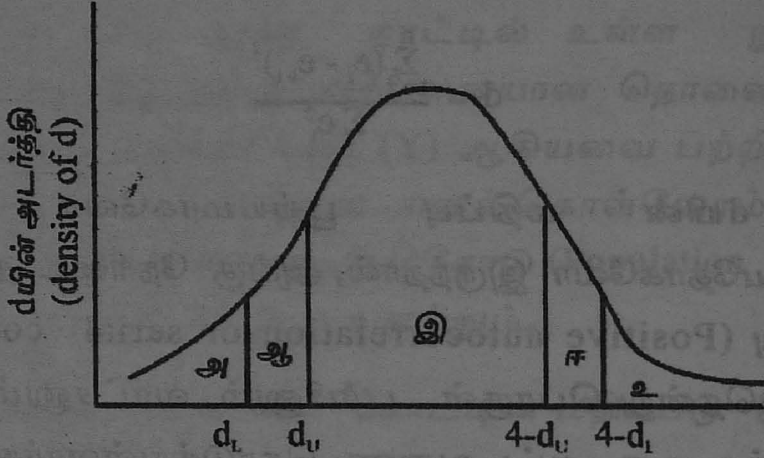
$d < d_L$ என்றிருந்தால், அங்கு நேரிடைத் தொடர் ஒட்டுறவு (Positive autocorrelation) உள்ளது என்று பொருள். (வரைபடம் 23இல் பகுதி அ)

$d > d_U$ என்றிருந்தால், அங்கு நேரிடைத் தொடர் ஒட்டுறவு இல்லை என்று பொருள். (பகுதி இ)

$d_L < d < d_U$ என்றிருந்தால் முடிவு செய்வது சிரமம் என்று பொருள். (பகுதி ஆ)

$d > (4-d_L)$ என்றிருந்தால் அங்கு எதிரிடைத் தொடர் ஒட்டுறவு என்று பொருள்.

வரைபடம் - 23



அ - நேரிடை தொடர் ஒட்டுறவு உள்ளது எனக் காட்டும் பகுதி

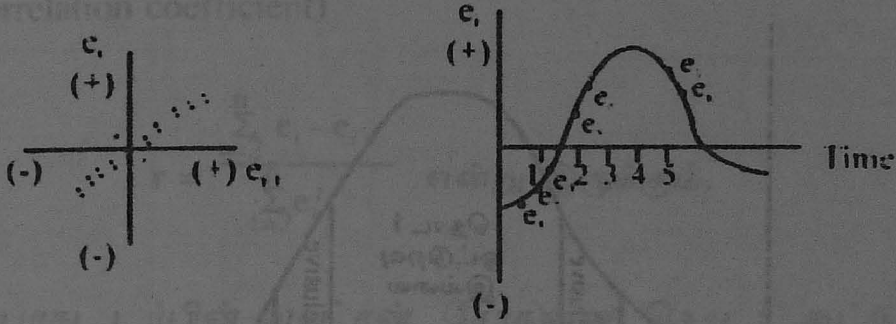
ஆ, ஈ - முடிவு செய்ய இயலாததைக் காட்டும் பகுதிகள்

இ - தன்னிச்சையாக உள்ளது / தொடர்பில்லை எனக் காட்டும் பகுதி

உ - எதிரிடை தொடர் ஒட்டுறவு (negative serial correlation) உள்ளது எனக் காட்டும் பகுதி.

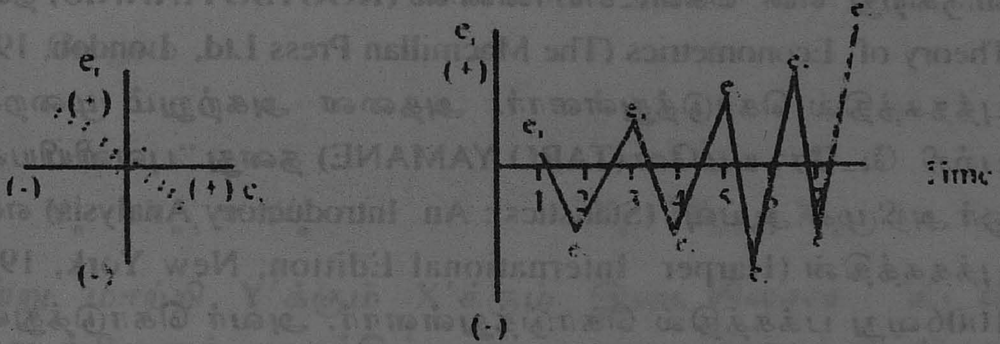
நாம் எடுத்துக்கொண்ட உதாரணத்தில் கணக்கிடப்பட்ட dன் (Calculated d value) மதிப்பு 0.9 என்றும் கொள்வோம். n=20 என்றும் இரண்டு தன்னிச்சையான மாறிகள் (independent variables) உள்ளன என்றும் கொள்வோம். ஐந்து விழுக்காடு முக்கியத்துவ அளவு (level of significance) கொண்டு டர்பின்-வாட்சன் அட்டவணையைப் பார்த்தால் அதில் $d_L = 1.10$ என்றும் $d_U = 1.54$ என்றும் இருக்கின்றன. எனவே $d < d_L$. எனவே, அங்கு நேரிடைத் தொடர் ஒட்டுறவு உள்ளதெனப் பொருள். அப்படியானால், வரைபடம் 24இல் உள்ளது போன்று e தன் பரவலைக் கொண்டிருக்கும்.

வரைபடம் - 24



வரைபடம் 25இல் உள்ளதுபோல் e தன் பரவலைக் கொண்டிருந்தால் அங்கு எதிரிடைத் தொடர் ஒட்டுறவு உள்ளதெனப் பொருள்.

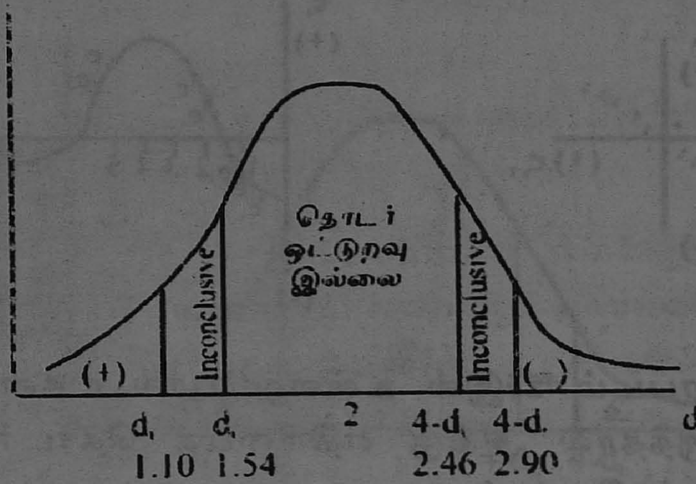
வரைபடம் - 25



எடுத்துக்கொள்ளப்பட்ட கணக்கில் d = 1.4 என்று வந்திருந்தால் முடிவு செய்ய முடியாத நிலை வந்திருக்கும். d = 1.7 என்று வந்திருந்தால் 'e'க்கள் தன்னிச்சையாக இயங்குகின்றன என்று சொல்லலாம். d = 3 என்று வந்திருந்தால் அங்கு எதிரிடைத் தொடர் ஒட்டுறவு உள்ளது என்று பொருள்.

தொடர் ஒட்டுறவு உள்ளது ஒரு பிரச்சனை. ஏனெனில் தொடர் ஒட்டுறவு இருக்கும்போது, பண்பலகுகளின் மாறுபாடு (variance) உண்மையில் உள்ளதைவிடக் குறைத்துக் காட்டப்படுகிறது. எனவே தொடர் ஒட்டுறவினை தவிர்க்க அல்லது அகற்றுவதற்கு வழி காண வேண்டும்.

வரைபடம் - 26



தொடர் ஒட்டுறவு என்னென்ன காரணங்களினால் வருகிறது என கௌட்சியான்னிஸ் (KOUTSOYIANNIS) தனது Theory of Econometrics (The Macmillan Press Ltd, London, 1981) புத்தகத்தில் கொடுத்துள்ளார். அதனை அகற்றும் முறைகள் பற்றி டேரோ யமனே (TARO YAMANE) தனது "புள்ளியியல் : ஓர் அறிமுக ஆய்வு" (Statistics : An Introductory Analysis) என்ற புத்தகத்தின் (Harper International Edition, New York, 1973) 1006வது பக்கத்தில் கொடுத்துள்ளார். அவர் கொடுத்துள்ள முறைகளில் (இடம் போதாமை காரணமாக) ஒன்று மட்டும் இங்கு தரப்படுகிறது.

தொடர் ஒட்டுறவினை அகற்றுவதற்கான வழி

தொடர் ஒட்டுறவினை அகற்றுவதற்கான வழிகளில் ஒன்று கிடைத்துள்ள புள்ளி விபரங்களைப் பொருத்தமாக மாற்றுவது (transforming the data) ஆகும். உதாரணமாக, உடன் தொடர்புப் போக்குக் கோடு

$$Y = a + bX$$ என இருந்தால்

தொடர் ஒட்டுறவுக் கெழுவின் மதிப்பீடு (estimated autocorrelation coefficient)

$$r = \frac{\sum_{t=2}^{n} e_t - e_{t-1}}{\sum_{t=2}^{n} e_{t-1}^2}$$ என்று இருக்கும்.

இப்போது டர்பின்-வாட்சன் சோதனை தொடர் ஒட்டுறவு இருக்கிறதென்று காட்டினால் கீழ்வருமாறு புள்ளிவிபரங்களை மாற்றி தொடர் ஒட்டுறவினை நீக்கிவிடலாம்.

$Y_2 - rY_1 = Y'_2$ $X_2 - rX_1 = X'_2$

$Y_3 - rY_2 = Y'_3$ $X_3 - rX_2 = X'_3$

மேலும்

....

$Y_n - rY_{n-1} = Y'_n$ $X_n - rX_{n-1} = X'_n$

என்று மாற்றி Y'க்கும் X'க்கும் இடையிலான உடன் தொடர்பினைக் கணக்கிட்டால், அங்கு தொடர் ஒட்டுறவு நீக்கப்பட்டு உடன்தொடர்புப் பண்பலகுகளின் மாறுபாடு உள்ளவாறே காட்டப்படும்.

கீழ்க்காணும் உத்தேசப் புள்ளி விபரங்களைக் கொண்டு தொடர் ஒட்டுறவுக் கெழுவின் மதிப்பைக் கண்டுபிடித்து, புள்ளி விபரங்களிலிருந்து தொடர் ஒட்டுறவை நீக்கும் முறையினைக் காணலாம்.

அட்டவணை - 25

தொடர் எண்.	Y	X	$\hat{Y}$	$e_t = Y - \hat{Y}$	$\Delta e_t = e_t - e_{t-1}$
1	33	10	34.15	-1.15	-
2	34	10	34.15	-0.15	1.00
3	38	11	36.90	1.10	1.25
4	43	12	39.65	3.35	2.25
5	46	12	39.65	6.35	3.00
6	46	13	42.40	3.60	2.75
7	45	13	42.40	2.60	-1.00
8	37	13	42.40	-5.40	-8.00
9	40	13	42.40	-2.40	3.00
10	38	13	42.40	-4.40	2.00
11	40	14	45.15	-5.15	0.75
12	43	14	45.15	-2.15	3.00
13	44	15	47.90	-3.90	-1.75
14	54	16	50.65	3.35	7.25
15	55	16	50.65	4.35	1.00
Sum	635	195		0	
Mean	42.4	13		0	

அட்டவணை 25இல் கொடுக்கப்பட்ட புள்ளி விபரங்களிலிருந்து

$$\hat{Y} = 6.65 + 2.75X$$

$$d = \frac{\sum_{2}^{15} (e_t - e_{t-1})^2}{\sum_{1}^{15} e_t^2}$$

$$\sum_{1}^{15} e_t^2 = 204.5991$$

$$\sum_{t=2}^{15} (e_t - e_{t-1})^2 = \sum_{t=2}^{15} (\Delta e_t)^2 = 167.357$$

$$d = 0.81806$$

டர்பின் - வாட்சன் பட்டியலிலிருந்து n 15ஆக இருக்கும்போது, தன்னியல்பான மாறி (independent variable) 1ஆக இருக்கும்போது 5 விழுக்காடு முக்கியத்துவத்தில் $d_L = 1.08$, du=1.36. எனவே, $d < d_L$. ஆகையால், அங்கு நேரிடைத் தொடர் ஒட்டுறவுப் பிரச்சனை உள்ளது. இத்தொடர் ஒட்டுறவுப் பிரச்சனையை நீக்குவதற்கு புள்ளி விபரங்களைச் சரிப்படுத்த வேண்டியுள்ளது.

$$Y'_t = Y_t - rY_{t-1} \quad t = 2, 3, ..., 15; \quad X'_t = X_t - rX_{t-1} \; t = 2, 3, ..., 15$$

இதில் r என்பது முதல்நிலை தொடர் உடன்கோட்டு முறைக்கெழு (coefficient of first order auto regressive scheme)

$$e_t = re_{t-1} + u_t$$

$$r = \frac{\sum_{2}^{15} e_t - e_{t-1}}{\sum_{2}^{15} e_{t-1}^2} = \frac{110.2899}{185.6774} = 0.5940$$

இந்த *0.5940*ஐப் பயன்படுத்தி X'யை Y'யையும் கணக்கிட்ட பின்பு,

$$\hat{Y}' = 0.4538 + 3.1972X'$$

$$\sum_{2}^{15} e' = 135.09946$$

$$\sum_{3}^{15} (\Delta e'_t) = 255.71191$$

$$d = \frac{255.71}{135.10} = 1.89$$

அட்டவணை - 26

வ.எண்.	Y′	X′	$\hat{Y}'$	e'_t	$\Delta e'_t$
1	-	-	-	-	-
2	14.4	4.1	13.4	0.9	-
3	17.8	5.1	16.6	1.2	0.2
4	20.4	5.5	17.9	2.5	1.3
5	20.4	4.9	16.0	4.4	1.9
6	18.7	5.9	19.2	-0.6	-5.0
7	17.7	5.3	17.3	0.3	0.9
8	10.3	5.3	17.3	-7.1	-7.4
9	18.0	5.3	17.3	0.7	7.8
10	14.3	5.3	17.3	-3.1	-3.8
11	17.4	6.3	20.5	-3.1	0.0
12	19.3	5.7	18.6	0.6	[illegible]
13	18.5	6.7	21.8	-3.4	-3.9
14	27.9	7.1	23.1	4.7	8.1
15	22.9	6.5	21.2	1.7	-3.0
Sum	257.9	78.7			
Mean	18.4	5.6			

டர்பின் - வாட்சன் பட்டியல்படி (n=15) $d_L = 1.08$ $d_U = 1.36$

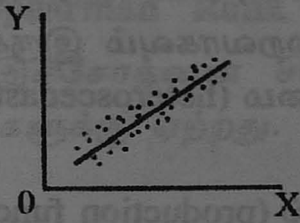
கணக்கிடப்பட்ட d (1.89) > d_U (1.36)

எனவே அங்கு நேரிடைத் தொடர் ஒட்டுறவு இல்லை. மேலும் கணக்கிடப்பட்ட d (1.89) < ($4-d_U$). எனவே, எதிரிடைத் தொடர் ஒட்டுறவும் இல்லை. ஆகவே, சரிபடுத்தப்பட்ட புள்ளி விபரங்களில் தொடர் ஒட்டுறவு (serial correlation or auto correlation) இல்லை என முடிவு செய்யலாம்.

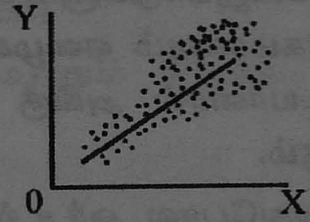
பன்முகத்தன்மை (Heteroscedasticity)

இயைபிலா மாறியான (random variable) 'u'ன் நிகழ்தகவுப் பரவல் (Probability distirbution) Xன் எல்லா மதிப்புக்களுக்கும் சமமாக இருக்கும்போதுதான் (Homoscedasticity) Y = a + bXன் பண்பலகுகளின் மாறுபாடுகள் (variance) பண்பலகுகளின் முக்கியத்துவத்தைச் சோதிக்கும்போது (tests of significance) பயனுள்ளவையாக இருக்கும். அவ்வாறின்றி, 'u'ன் நிகழ்தகவுப் பரவல் Xன் எல்லா மதிப்புக்களுக்கும் சமமாக இன்றி வேறுபட்டால், அந்தப் பண்பலகுகளின் முக்கியத்துவத்தைச் சேகரிக்கும்போது பயன்படுத்தப்படும் மாறுபாடுகளுக்கான சூத்திரங்கள் சரியில்லாமல் போய்விடும். இந்தப் பிரச்சனைக்கு பன்முகத்தன்மை (heteroscedasticity) என்று பெயர்.

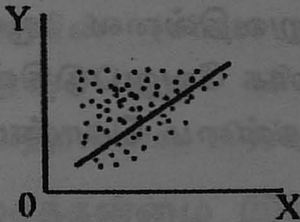
வரைபடம் - 27



(அ) ஒருமுகத்தன்மை
(Homoscedasticity)



(ஆ) கூடும் 'u'ன் மாறுபாடு
(Increasing vaiance of u)



(இ) குறையும் 'u'ன் மாறுபாடு
Decreasing variance of 'u'

அதாவது, var (u_i) = σ^2_u not constant என்று சொல்கிறோம். அப்படியானால் ஒவ்வொரு 'u'ன் மாறுபாடும் Xன் மதிப்பைப் பொறுத்து அமையும். அதாவது $\sigma^2_u = f(x)$. ஒருமுகத் தன்மையும் (homscedasticity) பன்முகத்தன்மையும் வரைபடம் 27இல் காட்டப்பட்டுள்ளன.

பொதுவாக, Xன் மதிப்பு கூடக்கூட σ^2_uன் மதிப்பும் கூடுவதாகக் கருதப்படுகிறது. அதாவது $\sigma^2_u = K^2X^2$ இங்கு 'K'ன் மதிப்பு மதிப்பிடப்படவுள்ளது.

பன்முகத்தன்மைக்கான காரணங்கள்

ஒரு மாறியில் காணப்படும் மாறுபாடுகள் கூடிக் கொண்டு போகும் நிலை பல சமயங்களில் காணப்படுகிறது. உதாரணத்திற்கு, பணக்காரர்களிடையே சேமிப்பில் அதிக வேறுபாடுகள் (variability) காணப்படலாம். அந்த அளவுக்கு ஏழைகளின் சேமிப்பில் வேறுபாடுகள் காணப்படாது. இந்நிலையில் வருமானம் (Y) கூடும்போது சேமிப்பில் (S) உள்ள வேறுபாடுகளும் கூடலாம். எனவே, பணக்காரர்களுக்கு 'u' அதிகமாகவும் ஏழைகளுக்கு 'u' குறைவாகவும் இருக்கும். அப்படியானால் அங்கு பன்முகத்தன்மை (heteroscedasticity) இருக்கும்.

அதுபோல ஓர் உற்பத்திச் சார்பில் (production function : $X = aL^{\alpha}K^{\beta}e^{u}$), தொழில் முயலும் தன்மை, தொழில் திறமை, தொழில் முறைமைகள், அமைப்பு சார்ந்த திறமைகளில் உள்ள வித்தியாசங்கள் 'u'ன் மதிப்பை நிர்ணயிக்கின்றன. இவை சிறிய நிறுவனங்களில் அதிகமாக மாறுவதில்லை. ஆனால், பெரிய நிறுவனங்களில் இவை அதிகமாக வேறுபடுகின்றன. அப்படியிருக்கும்போது, 'u' பன்முகத்தன்மை கொண்டதாக இருக்கும்.

பன்முகத்தன்மையின் விளைவுகள்

'u' பன்முகத்தன்மை உடையதாக இருந்தால், பண்பலகுகளின் மாறுபாட்டை (variance of a and variance of b)

மதிப்பிடும் சூத்திரங்கள் பயனில்லாததாகி விடும். ஆனாலும், இது பண்பலகுகளின் மதிப்பீடுகளைப் பாதிக்காது. இருப்பினும், இப்படிப்பட்ட பண்பலகுகளைக் கொண்டு கணிக்கும் கணிப்புக்கள் திறனற்றவைகளாகவே (inefficient) இருக்கும்.

பன்முகத்தன்மைக்கான சோதனைகள்

'u' ஒருமுகத்தன்மையுடையதாக இருக்கிறதா, பன்முகத் தன்மை உடையதாக இருக்கிறதா என்று அறிய பல சோதனைகள் உள்ளன. அவற்றில் ஒன்றுமட்டும் உதாரணத்திற்காக இங்கு கொடுக்கப்படுகிறது. மற்ற முறைகளைக் கற்றுக் கொள்ள விரும்புபவர்கள், கௌட்சியான்னிஸ் (KOUTSOYIANNIS) எழுதியுள்ள எகானாமெட்ரிக்ஸ் புத்தகத்தில் 185ஆவது பக்கத்திலிருந்து படித்துப் பார்க்கலாம்.

ஸ்பியர்மேன் தர ஒட்டுறவுச் சோதனை
(The Spearman Rank Correlation Test)

இச்சோதனை சிறிய (<30) மற்றும் பெரிய (30<) மாதிரிகளுக்கு ஏற்றது.

$$r_{e.x} = 1 - \frac{6\Sigma d^2}{(n^3 - n)}$$

d என்பது X மற்றும் e இணை (pair)களுக்கிடையேயுள்ள வித்தியாசம். n என்பது ஒரு மாதிரியில் (sample) உள்ள புள்ளிகளின் (observations) எண்ணிக்கை. இதில் தர ஒட்டுறவுக்கெழு மிக அதிகமாக இருந்தால் அங்கு பன்முகத்தன்மை பிரச்சனை உள்ளது என்று பொருள்.

பன்முகத்தன்மை இடையூறுகளை நீக்குதல்
(Solutions for Heteroscedastic Disturbances)

பன்முகத்தன்மை பிரச்சனை இருக்கிறதென்று கண்டுபிடிக்கப்பட்டால் அதை நீக்குவதற்கு முதலில்

எடுத்துக்கொண்ட சார்பினைப் (Model) பொருத்தமாக மாற்ற (transform) வேண்டும். உதாரணத்திற்கு $Y = a + b_1X_1 + u$ என்று சார்பு இருந்து, அதில் $\sigma^2_u = K^2X^2$ என்று இருந்தால்,

$$\frac{\sigma^2_u}{X^2} = K^2$$

முதலில் இருந்த சார்பினை (Model). X ($\sqrt{X^2}$) ஆல் வகுப்பதே பொருத்தமான மாற்றல் ஆகும். மாற்றலுக்குப் பின்னர்

$$\frac{Y}{X} = \frac{a}{X} + \frac{b_1X}{X} + \frac{u}{X}$$

$\frac{Y}{X} = \frac{a}{X} + b_1 + \frac{u}{X}$ என்று ஆகும். இதில் $\frac{u}{X}$ ஒருமுகத்தன்மை கொண்டதாக இருக்கும்.

இதற்கு எண்களைப் பயன்படுத்திய எடுத்துக்காட்டுகள் பார்க்க வேண்டுமானால் கௌட்ஸ்யான்னிஸ் எழுதியுள்ள எகானாமெட்ரிக்ஸ் புத்தகத்தின் *192* மற்றும் *193*ஆம் பக்கங்களைப் புரட்டலாம்.

பன்முக நேரிடைத்தன்மை (Multicollinearity)

முன்னரே கூறியதுபோல், பல மாறிகள் ஒன்றோடொன்று நெருக்கமான தொடர்பு கொண்டவைகளாக இருக்கும். ஒன்றில் ஏற்படும் மாற்றம் நேரிடையாகவும் முழுமையாகவும் மற்றொன்றையும் பாதிக்கும்; மற்றொன்றாலும் பாதிக்கப்படலாம். ஒருவர் தன் வருமானம் கூடக்கூட அவர் செலவையும் கூட்டிவந்தால், அவ்விரு மாறிகளுக்கும் இடையே ஒரு பூரண நேரிடை உறவு இருக்கலாம்; அப்பொழுது ஒட்டுறவுக்கெழு +1ஆக வரலாம். இந்த இரண்டு மாறிகளையும் ஒரு சார்பில் (function) சார்பிலா மாறிகளாகக் (independent variables : X_1X_2) கொண்டு அவை இரண்டிலும் ஏற்படுகின்ற மாறுபாடுகள் ஒரு சார்பு மாறியை

(Dependent variable:Y) பாதிக்கின்றன என ஒரு உறவுமுறையை (Model) ஏற்படுத்தினோமேயானால், அங்கு பன்முக நேரிடைத்தன்மை பிரச்சனை எழும்.

$Y = a + b_1X_1 + b_2X_2$ என்று கொண்டு X_1க்கும் X_2க்கும் இடையே ஒரு பூரண உறவு இருந்து X_1க்கும் X_2க்கும் இடையேயான ஒட்டுறவுக்கெழு +1ஆக இருந்தால், அங்கு பன்முக நேரிடைத்தன்மை இடையூறு உள்ளதென்று பொருள். இந்தப் பிரச்சனையைப் பற்றி டேரோ யமனே (TARO YAMANE) தன் புள்ளியியல் : ஓர் அறிமுக ஆய்வு (Statistics : An Introductory Analysis) என்ற புத்தகத்தின் மூன்றாம் பதிப்பில் (Third edition) பக்கம் 1009லிருந்து எழுதியுள்ளார்.

$$Y = a + b_1X_1 + b_2X_2 + e$$ என்பதில்

X_1 = தொழில்துறை உற்பத்திக் குறியீடாகவும்

X_2 = தேசிய மொத்த உற்பத்தியாகவும் இருந்தால் அவற்றிற்கிடையே பூரண நேரிடை ஒட்டுறவு (Multicollinear relationship) இருக்க வாய்ப்பிருக்கின்றதென்றும், X_1 = சூரிய ஒளியாகவும், X_2 = வெப்பநிலையாகவும் இருந்தால் அவற்றிற்கிடையேயும் பூரண நேரிடை ஒட்டுறவு இருக்க வாய்ப்பிருக்கின்றதென்றும் டேரோ யமனே தன் புத்தகத்தின் 1009வது பக்கத்தில் எழுதியுள்ளார்.

பன்முக நேரிடைத்தன்மையின் விளைவுகள் (Consequences of Multicollinearity)

$r_{x_1x_2} = 1$ ஆக இருந்தால், உடன் தொடர்புப் போக்கின் பண்பலகுகளின் மதிப்பீடுகள் (estimates of the regression coefficient) தீர்மானிக்கப்பட முடியாமல் (indeterminate) போகலாம்.

$\hat{b}_1$ ஐக் காண்பதற்கான சூத்திரம் : $$\frac{(\Sigma x_1y)(\Sigma x_2^2) - (\Sigma x_2y)(\Sigma x_1x}{(\Sigma x_1^2)(\Sigma x_2^2) - (\Sigma x_1x_2)^2}$$

$\hat{b}_2$ ஐக் காண்பதற்கான சூத்திரம்: $\dfrac{(\Sigma x_2 y)(\Sigma x_1^2) - (\Sigma x_1 y)(\Sigma x_1 x_2)}{(\Sigma x_1^2)(\Sigma x_2^2) - (\Sigma x_1 x_2)^2}$

இதில் x_2க்குப் பதிலாக ஒரு மாறிலியை (Constant : K) x_1 உடன் பெருக்கி (Kx_1) அதைப் பயன்படுத்தினால்,

$$\hat{b}_1 = \frac{K^2(\Sigma x_1 y)(\Sigma x_1^2) - K^2(\Sigma x_1 y)(\Sigma x_1^2)}{K^2(\Sigma x_1^2)^2 - K^2(\Sigma x_1^2)^2} = \frac{0}{0}$$

$$\hat{b}_2 = \frac{K(\Sigma x_1 y)(\Sigma x_1^2) - K(\Sigma x_1 y)(\Sigma x_1^2)}{K^2(\Sigma x_1^2)^2 - K^2(\Sigma x_1^2)^2} = \frac{0}{0}$$

இவ்வாறாகப் பண்பலகுகள் தீர்மானிக்கப்பட முடியாமல் போகின்றன.

அதுபோல், பண்பலகுகளின் திட்டப்பிழைகளும் (standard errors) அளவிட முடியாது பெரியதாக (infinitely large) இருக்கும். அப்படியிருந்தால், பண்பலகுகளின் முக்கியத்துவத்தைச் சரியாகச் சோதனை செய்து பார்க்க முடியாது. இதுபற்றி மேலும் விளக்கங்கள் வேண்டுவோர் டேரோ யமனேயின் புள்ளியியல் புத்தகத்தையோ கௌட்சியான்னிசின் எகானாமெட்ரிக்ஸ் (பக்கம் 234) புத்தகத்தையோ படிக்கலாம்.

அதிகமான சாரா மாறிகள் (independent variables)

சில நிகழ்வுகள் பல மாறிகளால் பாதிக்கப்படலாம். அப்படிப்பட்ட சூழ்நிலைகளில் இரண்டு பிரச்சனைகள் எழலாம். முதலாவதாக, ஒன்றுக்கும் மேற்பட்ட இணைகளுடன் உள்ள ஒட்டுறவுக்கெழு 1ஆக இருக்கலாம். அப்படியானால் அவைகளிடையே பன்முக நேரிடைத்தன்மை (Multicollinearity) பிரச்சனை இருக்கிறதென்று பொருள். அது பண்பலகுகளை மதிப்பிடுவதில் பிரச்சனையைத் தரும்.

இரண்டாவதாக, மாறிகள் புள்ளிகளைவிட அதிகமாகி விடலாம் (number of variables > number of observations). அப்படியிருந்தால் மதிப்பிட வேண்டிய பண்பலகுகள் அங்குள்ள சமன்பாடுகளை விட அதிகமாகிவிடும். அப்படி இருந்தாலும் பண்பலகுகளை மதிப்பிடுவது இயலாததாகி விடும்.

இந்த சமயங்களில் தலைமைக் கூறு ஆய்வு (PRINCIPAL COMPONENT ANALYSIS) பயன்படுகிறது. காரணி ஆய்வில் (FACTOR ANALYSIS) இது ஒரு சிறப்பான இடம்பெறுகிறது. இது பற்றிய அதிக விபரங்களுக்கு சி.ஆர்.கோத்தாரி (C.R.Kothari) எழுதியுள்ள ஆய்வு முறைகள் (Research Methodology) எனும் புத்தகத்தின் இரண்டாம் பதிப்பின் (II edition, Wishwa Prakashan Publication, New Delhi, 1995) 368ஆம் பக்கத்தைப் பார்க்கலாம். அல்லது, கௌட்சியான்னிஸ் எழுதியுள்ள எகானமெட்ரிக்ஸின் கோட்பாடு (Theory of Econometrics) என்ற புத்தகத்தின் 424ஆம் பக்கத்தை படிக்கலாம்.

மற்ற பிரச்சனைகள்

தொடர்புப் போக்குடன் தொடர்புள்ள இன்னும் சில பிரச்சனைகளும் உள்ளன. இடம்போதாததால் அவற்றை இங்கு விவரிக்க இயலவில்லை. அவை பற்றியும் தெரிந்து கொள்ள டேரோ யமனேயின் புள்ளியியல் புத்தகத்தையும் கௌட்சியான்னிசின் எகானாமேட்ரிக்ஸின் கோட்பாடு புத்தகத்தையும் வாசிக்கலாம்.

இரண்டு அல்லது இரண்டுக்கும் மேற்பட்ட பண்புகளுக்கிடையே (attributes) உள்ள உறவினைக்காண வேண்டுமானால் இரண்டு வழிகள் உள்ளன. முதலாவதாக, பண்புகளை எண்ணிக்கைக்குள் வருவதுபோல் (Proxy) கொண்டுவருவது. உதாரணத்திற்கு, ஒரு நடவடிக்கையை ஒருவருடைய மதநம்பிக்கை பாதிக்கலாம் என்று எடுகோள்

கொண்டால் அவருடைய மதநம்பிக்கைக்கு ஓர் எண் (Score, code, scale) கொடுத்தால் ஆய்வைச் சிறப்பாகச் செய்யலாம். ஆனால், ஒரு பண்பு மாறிக்கு (attribute) எண் கொடுக்கும்போது மிகக் கவனமாகக் கொடுக்க வேண்டும். உதாரணத்திற்கு, இந்து மத நம்பிக்கைக்கு எந்த எண்ணைக் கொடுக்கலாம்? இஸ்லாம் மத நம்பிக்கைக்கு எந்த எண்ணைக் கொடுக்கலாம்? இது பற்றி முடிவு செய்வதற்கும் முன்பு ஆழ்ந்த ஆலோசனை தேவை. அதேபோல, ஒருவரின் படிப்பறிவை எப்படி எண்ணால் அளப்பது? அவர் எத்தனை ஆண்டுகள் பள்ளி / கல்லூரிக்குச் சென்றுள்ளார் என்பதை மட்டும் வைத்து, ஒருவரின் படிப்பறிவை எண்ணாலாக்கினால் அது எந்த அளவுக்குச் சரியாக இருக்கும்?

இவ்வாறு முடிவு செய்ய இயலாத சூழ்நிலைகளில், χ^2 சோதனையைப் பயன்படுத்தியும் இரண்டு அல்லது இரண்டுக்கும் மேற்பட்ட பண்புகளுக்கிடையேயுள்ள ஒட்டுறவினை எண்ணால் கொடுக்க முடியும். இந்த χ^2ஐ பின்னால் தெரிந்து கொள்ளலாம். ஆனால் இப்போது பண்புகளின் கூட்டுறவு (Association of Attributes) பற்றி மட்டும் இங்கு காணலாம்.

பண்புகளின் கூட்டுறவு (Association of Attributes)

கொடுக்கப்பட்டுள்ள மாறிகள் அளவின மாறிகளாக (Quantitative variables) இருந்தால், அவற்றிற்கிடையே உள்ள தொடர்பை ஆராய்வதற்கு உடன்தொடர்பு மற்றும் ஒட்டுறவுக்கெழுவினைப் பயன்படுத்தலாம் என இதுவரை விளக்கப்பட்டது. மாறாக, கொடுக்கப்பட்டுள்ள மாறிகள் பண்பின மாறிகளாக இருந்தால் (Qualitative variables) அவற்றிற்கிடையே உள்ள தொடர்பை ஆராய்வதற்கு பண்புகளின் கூட்டுறவு பயன்படுத்தப்படுகிறது. பண்பின மாறிகளாக கண்ணின் நிறம், மதம், சாதி, குணம், நேர்மை, கல்வியறிவு போன்றவற்றைக் கூறலாம். உதாரணத்திற்கு

மதத்திற்கும் கல்வியறிவிற்கும் இடையே உறவு உள்ளதா, அப்படியிருந்தால், அது நேரிடை உறவா, எதிரிடை உறவா என்று அறிய வேண்டிய அவசியம் வரலாம். அப்படிப்பட்ட சூழ்நிலைகளில் பண்புகளின் கூட்டுறவு காணும் முறை பயன்படும். பண்புகளை (A, B, C, ...) என்றும் அந்தப் பண்புகள் கொண்ட மனிதர்களை முறையே (A), (B), (C) ... என்றும் குறிக்கலாம். உதாரணத்திற்கு பெண்களை A என்று கொண்டால் பெண்களின் எண்ணிக்கையை (frequency) (A) என்று குறிக்கலாம். ஆண்களை α (ஆல்ஃபா) என்ற கிரேக்க எழுத்தால் குறிக்கலாம். ஆண்களின் எண்ணிக்கையை (α) (ஆல்ஃபா) என்று அடைப்புக்குறிக்கள் எழுதலாம். படித்தவர்களை B என்று கொண்டால் படித்தவர்களின் எண்ணிக்கையை (B) என்றும் படிக்காதவர்களை β (beta: பீட்டா) என்றும் படிக்காதவர்களின் எண்ணிக்கையை (β) என்று அடைப்புக்குறிக்குள்ளும் குறிக்கலாம். இந்த செய்தியைப் பயன்படுத்தி ஓர் இணைப்பட்டியலைக் கீழ்க்காணுமாறு உருவாக்கலாம்.

அட்டவணை - 27
(இணைப்பட்டியல் : Contingency Table)

	பெண்கள் A	ஆண்கள் α	மொத்தம்
படித்தவர்கள் B	(AB)	(αB)	(B)
படிக்காதவர்கள் β	(Aβ)	(αβ)	(β)
மொத்தம்	(A)	(α)	N

மூன்றாவதாக ஒரு பண்பு இருந்தால் C என்றும் அப்பண்பு இல்லாதவரை γ (gama : காமா) என்றும் கூறலாம். நான்காவதாக ஒரு பண்பு இருந்தால் அதனை D என்றும் அப்பண்பு இல்லாமையை δ (delta : டெல்டா) என்றும்

குறிக்கலாம். மொத்த எண்ணிக்கையைக் குறிக்க N பயன்படுத்தப்படுகிறது.

இந்த இணைப்பட்டியலில் (அட்டவணை 27) உள்ள சில கட்டங்கள் அல்லது அறைகளில் (cell) எண்கள் கொடுக்கப்பட்டிருந்தால் மீதமுள்ள எண்களை கொடுக்கப்பட்டுள்ள எண்களிலிருந்து தருவிக்க முடியும். உதாரணத்திற்கு அட்டவணை 28ஐப் பார்க்கலாம்.

அட்டவணை - 28

	பெண்கள் A	ஆண்கள் α	மொத்தம்
படித்தவர்கள் B	(AB) 15	(αB)	40
படிக்காதவர்கள் β	(Aβ)	(αβ)	
மொத்தம்	**50**	(α)	**100**

இந்த அட்டவணையில் (28) நான்கு எண்களே கொடுக்கப்பட்டுள்ளன. இன்னும் ஐந்து கட்டங்கள் (Cells) எண்களால் நிரப்பப்பட வேண்டும்.

(A) = 50 (AB) = 15 (B) = 40 N = 100

இந்த எண்களில் இருந்து αBயைக் கண்டுபிடிக்கலாம். அது 25. அதுபோல் Aβ = 35; αβ = 25, β = 60, α = 50.

பண்புகளின் உறவு

$\frac{(AB)}{(B)} = \frac{(A\beta)}{(\beta)} = \frac{(A)}{N}$ என இருந்தால் Aயும் Bயும் சார்பிலாப் பண்புகள் ஆகும். அப்படியானால் αஉம் βஉம் தனித்த (சார்பிலா) பண்புகளாகவே இருக்கும். கிடைத்த விபரங்களிலிருந்து கீழ்க்காணும் அட்டவணையைப் (29) பெறலாம்.

அட்டவணை - 29

	பெண்கள் A	ஆண்கள் α	மொத்தம்
படித்தவர்கள் B	15	25	40
படிக்காதவர்கள் β	35	25	60
மொத்தம்	**50**	**50**	**100**

இதில் பாலினமும் படிப்பறிவும் சார்புள்ளவையா, சார்பற்றவையா என்று கண்டுபிடிக்கலாம்.

$$\frac{(AB)}{(B)} = \frac{15}{40}; \quad \frac{(A\beta)}{(\beta)} = \frac{35}{60}; \quad \frac{(A)}{N} = \frac{50}{100}$$

$$(AB) = 15; \quad \frac{(A)(B)}{N} = \frac{(50)(40)}{100}$$

இவற்றிலிருந்து A யும் B யும் சார்புள்ள பண்புகள் எனத் தெரியவருகிறது. அதாவது, கிடைத்த அலைவெண்ணும் (frequency) எதிர்பார்க்கப்படும் அலைவெண்ணும் (expected frequency) சமமாக இருந்தால் அப்பண்புகள் சார்பிலாப் பண்புகள் ஆகும். அவையிரண்டும் சமமாக இல்லையெனில் சார்புள்ள பண்புகளாகும். அட்டவணை 29ல் $(AB) = 15$; ஆனால், $\frac{(A)(B)}{N} = \frac{50 \times 40}{100} = 20$. எனவே $(AB) \neq \frac{(A)(B)}{N}$. எனவே A யும் B யும் சார்புடைய பண்புகள் ஆகும். அவ்வாறின்றி அட்டவணை 30இல் உள்ளதுபோல் இருந்தால் A யும் B யும் சார்பிலாப் பண்புகளாக இருக்கும்.

அட்டவணை - 30

	பெண்கள் A	ஆண்கள் α	மொத்தம்
படித்தவர்கள் B	20	40	60
படிக்காதவர்கள் β	30	60	90
மொத்தம்	**50**	**100**	**150**

அட்டவணை 30இல் கிடைத்த அலைவெண்ணும் [(AB) = 20] எதிர்பார்க்கப்படும் அலைவெண்ணும் $\left[\frac{(A)(B)}{N} = \frac{50 \times 60}{150} = 20\right]$ சமமாக இருக்கின்றன. அவைகளுக்கிடையே உள்ள வித்தியாசம் 0 (பூஜ்யம்). எனவே அவை சார்பிலாப் பண்புகள் ஆகும். இதே போன்ற கருத்தின் அடிப்படையில் χ^2ம் இருப்பதைப் பின்னர் அறியலாம்.

உறவின் திசைகள் (Directions of relationship)

கிடைத்த அலைவெண்ணுக்கும் எதிர்பார்க்கப்படும் அலைவெண்ணுக்கும் உள்ள வித்தியாசம் பூஜ்யமாக இருந்தால் எடுத்துக்கொண்ட பண்புகளிடையே உறவில்லையென்றும் அவ்விரண்டுக்கும் இடையே வித்தியாசம் இருந்தால் அந்தப் பண்புகளிடையே உறவு உள்ளதென்றும் அறிந்தோம். உறவு உள்ளதென்றால், நேரிடை (Positive) உறவா எதிரிடை (Negative) உறவா என்று எப்படி அறிவதென்பது பற்றி இனிப் பார்க்கலாம்.

$(AB) > \frac{(A)(B)}{N}$ என்றிருந்தால் $(AB) - \left[\frac{(A)(B)}{N}\right]$ ன் விடை +ல் வரும். அப்படியானால், அந்தப் பண்புகளிடையே உள்ள உறவும் நேரிடை உறவாகும். மாறாக $(AB) < \frac{(A)(B)}{N}$

என்றிருந்தால் (AB) − $\left[\frac{(A)(B)}{N}\right]$ ன் விடை −ல் வரும். அப்படியானால், அந்தப் பண்புகளிடையே உள்ள உறவும் எதிரிடை உறவாகும்.

கீழே தரப்படுகின்ற உதாரணத்தில் படிப்பறிவுக்கும் நேர்மையாக இருப்பதற்கும் எப்படிப்பட்ட தொடர்பு உள்ளது என்று கண்டுபிடிக்கலாம். 100 நபர்களைக் கொண்ட ஒரு மாதிரியில் 70 பேர் படிப்பறிவு உள்ளவர்கள்; 50 பேர் நேர்மையானவர்கள். படிப்பறிவு உள்ளவர்களில் 30 பேர் நேர்மையானவர்கள்.

அட்டவணை - 31

	படிப்பறிவு உள்ளவர்கள் A	படிப்பறிவு இல்லாதவர்கள் α	மொத்தம்
நேர்மையாக (B) வாழ்பவர்கள்	(30)	20	(50)
நேர்மையாக வாழாதவர்கள் (β)	40	10	50
மொத்தம்	(70)	30	(100)

இதில் (AB) = 30; $\frac{(A)(B)}{N} = \frac{70 \times 50}{100} = 35$

இதில் $\left[(AB) - \frac{(A)(B)}{N}\right] = (30-35) = -5$ ஆக இருப்பதால், கல்வியறிவு, நேர்மை ஆகிய பண்புகளிடையே எதிரிடை உறவு உள்ளது என்று பொருள்.

கூட்டுறவுக்கெழு (Coefficient of Association)

யூல் (YULE) என்பவர் கூட்டுறவுக் கெழுவினைக் காண கீழ்வரும் சூத்திரம் தந்துள்ளார்.

$$Q = \frac{(AB)\,(\alpha\beta) - (A\beta)\,(\alpha B)}{(AB)\,(\alpha\beta) + (A\beta)\,(\alpha B)}$$

A,B என்னும் இரு பண்புகளுக்கிடையே 'முற்றிலும் நேரிடை உறவு' (Complete Positive Association : Q = +1) இருந்தால் A பண்பு உடையவர்கள் B பண்பையும் உடையவர்களாகவும், A பண்பு இல்லாதவர்கள் B பண்பு இல்லாதவர்களாகவும் இருப்பார்கள்.

A,B என்னும் இரு பண்புகளிடையே 'முற்றிலும் எதிரிடை உறவு' (Complete Disassociation : Q = -1) இருக்கும்பொழுது, A பண்பு உடையவர்களெல்லாம் B பண்பு இல்லாதவர்களாகவும், B பண்பு உடையவர்களெல்லாம் A பண்பு இல்லாதவர்களாகவும் இருப்பார்கள்.

Q = 0 என்று இருந்தால் Aக்கும் Bக்கும் இடையே உறவில்லை; அவ்விரண்டும் தனித்த சார்பிலா பண்புகள் என்று பொருள். அட்டவணை 32இல் தரப்பட்டுள்ள விவரங்களுக்குக் கூட்டுறவு கெழு (Q) கணக்கிடலாம்.

அட்டவணை - 32

	அமைதியான வாழ்வு வாழ்பவர்கள் (A)	அமைதியான வாழ்வு பெறாதவர்கள் (α)	மொத்தம்
படிப்பறிவு உள்ளவர்கள் (B)	48	32	80
படிப்பறிவு இல்லாதவர்கள்(β)	12	8	20
மொத்தம்	**60**	**40**	**100**

$$Q = \frac{(AB)(\alpha\beta) - (A\beta)(\alpha B)}{(AB)(\alpha\beta) - (A\beta)(\alpha B)} = \frac{(48)(8) - (12)(32)}{(48)(8) + (12)(32)}$$

$$= \frac{(384) - (384)}{(384) + (384)} = \frac{0}{678} = 0$$

எனவே, Aயும் Bயும் தனித்த சார்பிலா பண்புகள் என்று பொருள்.

பண்புகளின் உறவினை அறிய χ^2 கை வர்க்கம் அல்லது சை வர்க்கம் பயன்படுகிறது. அதனைப் பிறகு காணலாம்.

7. நிகழ்தகவு
(PROBABILITY)

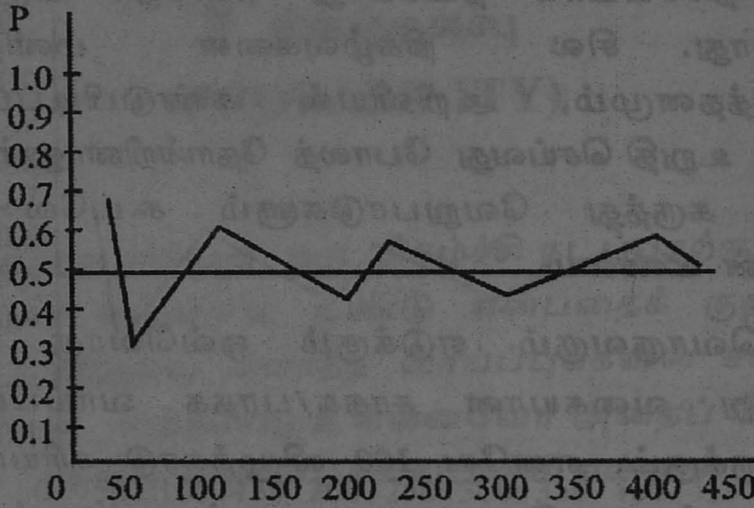
நிகழ்தகவு என்பது ஒரு நிகழ்ச்சி நடப்பதற்குத் தகுந்த வாய்ப்புக்கள் எவ்வளவு உண்டு என்பதைக் குறிக்கிறது. எனவே, நிகழ்தகவு மொத்த வாய்ப்புக்களில் சாதகமான வாய்ப்புக்கள் எவ்வளவு உள்ளனவோ அதைப் பொறுத்து அமையும். நிச்சயம் ஒரு நிகழ்ச்சி நடக்கும் என்றால் நிகழ்தகவு *1* என்றும் நிச்சயம் நடக்காது என்றால் நிகழ்தகவு *0* (பூஜ்யம்) என்றும் குறிக்கப்படுகிறது. சில சமயங்களில், முன்னரே பெற்ற அனுபவத்தைக்கொண்டு சாதகமான வாய்ப்புக்கள் மொத்த வாய்ப்புக்களில் எந்த அளவுக்கு உள்ளன என்று சொல்லமுடியலாம். ஆனால், நடைமுறையில் பல சமயங்களில் சாதகமான வாய்ப்பு எவ்வளவு, பாதகமான வாய்ப்பு எவ்வளவு என்று கூறமுடியாது. ஒரு வண்டி சாலையில் போகும்போது அது விபத்துக்குள்ளாகுமா ஆகாதா என்று கூறுவது அவ்வளவு எளிதல்ல; அது பல காரணிகளைப் பொறுத்து அமையும். ஆனால் சாலையில் போகும் *100* வண்டிகளில் ஒரு வண்டி விபத்துக்குள்ளாகும் என்று முன்னரே கூறிவிட்டால் ஒவ்வொரு வண்டியும் விபத்துக்குள்ளாவதற்கான வாய்ப்பு 1 ÷ 100 (0.01) என்று சொல்ல முடியும். ஆனால், இந்த நிகழ்தகவான 1 ÷ 100 அல்லது 0.01 என்பது எல்லா வண்டிகளுக்கும் சமமாக இருக்கும் என்று கூற முடியாது. கவனமாக ஓட்டப்படும் வண்டிகள் விபத்துக்குள்ளாக *0.01*ஐவிடக் குறைவான வாய்ப்பையும்; கவனக்குறைவாக ஓட்டப்படும் வண்டிகள் விபத்துக்குள்ளாகும் வாய்ப்பை *0.01*ஐ விட அதிகமாகவும் கொண்டிருக்கும். அன்றாடம் மனிதர்கள் எடுக்கும் எல்லா முடிவுகளுமே நிகழ்தகவோடு தொடர்பு

கொண்டனவாகவே உள்ளன. எதுவுமே நிச்சயம் நடக்கும் என்றோ நிச்சயமாக நடக்காது என்றோ உறுதியாகக் கூறமுடியாது. சில நிகழ்வுகளை மனிதர்களின் புத்திசாலித்தனமும், அறிவியல் கண்டுபிடிப்புக்களும் ஓரளவுக்கு உறுதி செய்வது போலத் தோன்றினாலும், இதில் பலவிதக் கருத்து வேறுபாடுகளும் கூடவே உள்ளன என்பதுதான் உண்மை.

ஒவ்வொருவரும் எடுக்கும் ஒவ்வொரு முடிவும் வெவ்வேறு வகையான சாதக/பாதக வாய்ப்புக்களை கொண்டிருக்கும். எனவே 100 விழுக்காடு சரியாக எந்த முடிவினையும் கூற இயலாது. ஆனாலும், ஆய்வுகள் முழுக்க முழுக்கச் சரியாகச் செய்யப்பட்டால், அந்த ஆய்வுகள் கூறுகின்ற முடிவுகள் சரியாக இருப்பதற்கான வாய்ப்புக்கள் அதிகம் என்று கருதப்படுவதால்தான் இன்று ஆய்வுகள் அதிகமாகிக் கொண்டிருக்கின்றன. ஆய்வுகள் செய்வதற்கான நிதி ஆதாரங்களும் கூடிக்கொண்டு இருக்கின்றன.

அனேகமாக, எல்லாவகையான மனித நடவடிக்கைகளும், மிகவும் குறிப்பாக, வணிக நடவடிக்கைகளும் நிச்சயமின்மையைக் (uncertainty) கொண்டுள்ளன. இச்சூழலில், ஏதேனும் ஒரு வழியில் நிச்சயமின்மையை அளவிட முடியும் என்றால், அது மிகவும் பலனுடையதாக இருக்கும். இதன் அடிப்படையில்தான், நிகழ்தகவும், நிகழ்தகவுப் பரவல்களும் (Probability distribution) கோட்பாட்டுப் பரவல்களும் (Theoretical distribution) கற்று அறியப்படுகின்றன.

வரைபடம் - 28



குறிப்பு : $P = \frac{\text{தலைகளின் எண்ணிக்கை}}{\text{நாணயங்களின் எண்ணிக்கை}}$ சுண்டப்படும் நாணயங்களின் எண்ணிக்கை கூடக்கூட, தலை தோன்றுவதற்கான நிகழ்தகவு 0.5ஐ நெருங்கி வருகிறது.

ஒரு நல்ல நாணயத்தைச் (unbiased coin) சுண்டிவிட்டு அதில் தலை (head) அல்லது பூ (tail) வருவதற்கான வாய்ப்பினைக் கூறுவது எளிதாக இருக்கலாம். அவ்வாறு கூறுவது சரியாக இருப்பதற்கும் வாய்ப்புக்கள் அதிகமாக இருக்கலாம். அதேபோல பகடைக்கட்டையிலும், விளையாடுதற்குப் பயன்படும் சீட்டுக்கட்டுகளிலும், குழந்தை பிறப்பிலும் சில கேள்விகள் கேட்டால், சரியான பதில் சொல்வது ஓரளவுக்கு எளிது. ஏனெனில், மேலே சொன்ன எடுத்துக்காட்டுகளில் சில செய்திகள் முன்னரே (apriori) அறியப்பட்டவை. ஒரு நல்ல நாணயத்தைச் சுண்டி (trial) விட்டால், அதன் மேற்பக்கம் தலை தோன்றுவதற்கான (event) வாய்ப்பினை ½ அல்லது 0.5 என்று சொல்லலாம். குறைவான எண்ணிக்கையில் நாணயங்களைச் சுண்டும்போது சரி பாதியளவு நாணயங்களில் தலையும் மீதி பாதியளவு

நாணயங்களில் பூவும் தோன்றாவிட்டாலும் ஓரளவுக்கு சரிபாதிக்குப் பக்கத்தில் இருக்கும் (வரைபடம் 28ஐப் பார்க்கலாம்). நாணயங்களின் எண்ணிக்கையைக் கூட்டக் கூட்ட தலை தோன்றுவதற்கான வாய்ப்பும் பூ தோன்றுவதற்கான வாய்ப்பும் நெருங்கி வந்து இரண்டு நிகழ்வுகளுக்குமான வாய்ப்பு 0.5 மற்றும் 0.5 ஆக முடியும். ஆனால் இந்தளவுக்குத் தோராயமாகக் கூட அன்றாடம் நடைபெறும் நிகழ்வுகளில் கூற முடியாமல் போகலாம். அப்பொழுது ஏற்கனவே கிடைத்துள்ள அலைவெண்களைக் கொண்டு ஓரளவுக்கு வாய்ப்புகள் பற்றிய பதிலைத் தர முயற்சிக்கலாம். உதாரணத்திற்கு ஒரு வகுப்பறையில் உள்ள 50 மாணவர்களில் 30 பேர் தேர்ச்சி பெற்றுள்ளார்கள் என்று கொண்டோமேயானால், அந்த வகுப்பறையிலிருந்து ஏதேனும் ஒரு மாணவரை அழைக்கும்போது வருபவர் தேர்ச்சி பெற்ற மாணவராக இருப்பதற்கான வாய்ப்பு 30 ÷ 50 (0.6) ஆகும்.

ஓர் ஆயுள் காப்பீட்டுக் கழகத்தில் பணம் கட்டி அங்கத்தினராகச் சேரும்போது, எத்தனை பேர் எந்த அளவு இழப்பினைச் சந்திக்க நேரிடும் என்று சொல்வது சற்றுச் சிரமமான செயல்தான். இருப்பினும், அனுபவத்தைக் கொண்டு ஓரளவுக்கு சரியான பதில் கூற முடியும். அவ்வாறாகத் தானே, ஆயுள் காப்பீட்டுக் கழகங்களின் பணிகள் தொடர்ந்து நடைபெற்று வருகின்றன. இதில் எதிர்பார்க்கும் இழப்பைவிட, சில எதிர்பாராத காரணங்களினால் உண்மையான இழப்பு அதிகமாகும்போது பிரச்சனைகள் வந்து விடலாம். எனவே, இம்மாதிரிச் சூழ்நிலைகளில் வெவ்வேறு நிகழ்வுகளின் வாய்ப்புக்களை முன்னரே சரியாகக் கூறும் திறமை வளர்ந்துவிட்டால் அதன் மூலம் பலரும் பலன் பெறலாம்.

நிகழ்வுகளின் வகைகள்

ஒவ்வொரு செயலிலும் (Process, experiment or trial) பல நிகழ்ச்சிகள் (events) இருக்கலாம். உதாரணத்திற்கு ஓர் ஊருக்குச் செல்வதென்பது செயல். அதில் புறப்படுவது, பேருந்தை அடைவது, நினைத்த ஊரை அடைவது போன்றவை நிகழ்ச்சிகளாகும். போகிறேன் என்பதை ஒரு செயலாகவும் புறப்படுகிறேன் என்பதை ஒரு நிகழ்வாகவும் கொள்ளலாம். காலை *10.00* மணிக்குப் புறப்பட்டேன்; *10* மணி *30* நிமிடத்திற்குப் பேருந்தை அடைந்தேன்; *12* மணிக்கு ஊரை அடைந்தேன். செல்கிறேன், போகிறேன் என்பன தொடர் நிகழ்வுகள்; ஒரு கால அளவைக் கொண்டிருக்கும். ஒரு பகடைக் கட்டையை உருட்டுவது, ஒரு நாணயத்தைச் சுண்டி விடுவது போன்றவைகளும் செய்கைகள் அல்லது சோதனைகள் எனலாம். அவற்றிலிருந்து கிடைப்பவைகளை விளைவுகள் எனலாம். அல்லது நிகழ்ச்சிகள் எனலாம். இந்த நிகழ்ச்சிகள் எவ்வாறு உடன் தொடர்புள்ள நிகழ்ச்சிகளைப் பாதிக்கின்றன அல்லது உடன் தொடர்புள்ள நிகழ்ச்சிகளால் பாதிக்கப்படுகின்றன என்பதை வைத்து அந்த நிகழ்ச்சிகளை வகைப்படுத்தலாம். நிகழ்தகவினைக் கணிக்கும்போது பயன்படுத்தப்படும் சூத்திரம் நிகழ்ச்சிகளின் வகையைப் பொறுத்து அமையும்.

பூரண நிகழ்ச்சிகள் (EXHAUSTIVE EVENTS)

ஒரு நாணயத்தைச் சுண்டி விடுவது சோதனை (experiment or trial) எனப்படுகிறது. அந்த நாணயத்தின் மேற்பக்கம் பூவோ தலையோ தோன்றலாம். ஒரு நல்ல நாணயத்திற்கு இரண்டு பக்கங்கள் உள்ளன. அதன் ஒரு பக்கத்தை பூ என்றும் மற்றொரு பக்கத்தை தலை என்றும் சொல்கிறோம். இந்த இரண்டு பக்கங்களைத் தவிர வேறு ஏதும் இல்லை; இரண்டு பக்கங்களும் ஒரே சமயம் மேலே தோன்ற முடியாது; பூ மற்றும் தலை தோன்றுவதற்கு (½) சமமான வாய்ப்பே

உள்ளது. இப்படிச் சூழ்நிலைகளைக் கொண்ட ஒரு சோதனையில், இரண்டு நிகழ்ச்சிகளுக்குமான வாய்ப்புக்களைக் கூட்டினால் ($\frac{1}{2} + \frac{1}{2} = 1$) ஒன்று. இந்த இரண்டு நிகழ்ச்சிகளுமே மொத்த வாய்ப்பான ஒன்றினை பூர்த்தி செய்து விடுவதாலும் வேறு எந்த வாய்ப்புக்களுக்கும் இடம் இல்லாததாலும் அவ்விரு நிகழ்ச்சிகளையும் பூரண நிகழ்ச்சிகள் எனலாம். ஒரு பகடைக்கட்டை (die)யில் ஆறு பக்கங்கள் உள்ளன. அந்த ஆறு பக்கங்களும் 1 முதல் 6 வரை எண்கள் பொறிக்கப் பட்டிருந்தால், அந்த ஆறு நிகழ்ச்சிகளின் மொத்த நிகழ்தகவும் ($\frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6}$) ஒன்று (1). எனவே, அந்த ஆறு நிகழ்வுகளும் பூரண நிகழ்வுகள் எனலாம்.

சமவாய்ப்பு நிகழ்ச்சிகள்
(Equiprobable events or equal likely events)

ஒரு நாணயத்தைச் சுண்டும்போது, நாணயத்தின் மேற்பக்கம் தலையாகவோ பூவாகவோ இருக்கும்படி நாணயம் விழலாம். இரண்டும் ஒவ்வொரு பக்கமாக இருப்பதாலும் நாணயத்தில் இரண்டே பக்கங்கள் இருப்பதாலும், பூ தோன்றுவதற்கு எவ்வளவு வாய்ப்போ அதே அளவு வாய்ப்புதான் தலை தோன்றுவதற்கும் உள்ளது. இரண்டுக்கும் சமவாய்ப்புக்கள் உள்ளதால் அவற்றை (பூவையும் தலையையும்) சமவாய்ப்பு நிகழ்ச்சிகள் எனலாம். அதுபோல, ஆறு பக்கங்கள் கொண்ட பகடைக் கட்டையில் 1, 2, 3, 4, 5, 6 என்று ஒவ்வொரு பக்கத்திற்கும் ஒன்றாகக் குறிப்பிட்டு இருந்தால் அந்த ஆறு நிகழ்வுகளுக்கும் சம வாய்ப்புக்களே உள்ளன. எனவே அவற்றைச் சமவாய்ப்பு நிகழ்ச்சிகள் எனலாம்.

ஒன்றையொன்று விலக்கிடும் நிகழ்ச்சிகள்
(Mutually exclusive events)

ஒரு நாணயத்தைச் சுண்டிவிட்டால் நாணயத்தின் மேலே தோன்றுவது தலையாக இருக்கும் அல்லது பூவாக இருக்கும்

இரண்டும் (நல்ல நாணயமாக இருந்தால்) மேலே தோன்ற வாய்ப்பில்லை. பூ தோன்றினால் தலை வராது. மேலே தலை வந்தால் பூ வராது. ஒன்று நிகழ்வது மற்றொன்று நிகழ்வதை விலக்கி விடுவதால் அவை ஒன்றையொன்று விலக்கிடும் நிகழ்ச்சிகள் எனப்படுகின்றன.

அதுபோல, ஒரு பக்கத்திற்கு ஓர் எண்ணாக ஆறு பக்கங்களிலும் 1, 2, 3, 4, 5, 6 ஆகிய எண்களைத் தாங்கிய பகடைக் கட்டையை உருட்டும்போது, அந்த ஆறு எண்களில் ஒரு சமயத்தில் ஒரு எண்தான் தோன்ற முடியும்; மீதமுள்ள ஐந்து எண்களும் விலக்கப்படுகின்றன. எனவே, இந்த நிகழ்வுகளும் ஒன்றையொன்று விலக்கும் நிகழ்வுகளே.

சார்புடைய நிகழ்ச்சிகள் (Dependent events)

இதற்கு முன்னர் கூறப்பட்டது ஒன்றையொன்று விலக்கும் நிகழ்ச்சி. இந்நிகழ்ச்சிகளும் ஒன்றை ஒன்று பாதிக்கின்றன. சார்புடைய நிகழ்ச்சிகள் ஒன்றையொன்று பாதிக்கும்; ஆனால் விலக்காது. உதாரணத்திற்கு ஒரு பெட்டியில் 4 வெள்ளைப் பந்துகளும் 5 பச்சைப் பந்துகளும் இருக்கிறதென்று கொள்வோம். அவற்றில் முதலில் ஒரு பந்து எடுக்கப்பட்டு, அது பெட்டிக்குள் போடப்படாமல், மறுபடியும் அந்தப் பெட்டிக்குள் இருந்து இன்னொருமுறை ஒரு பந்து எடுத்தால் அந்தப்பந்து பச்சைப் பந்தாக இருப்பதற்கு என்ன நிகழ்தகவு என்று கேட்டால், இதில் இரண்டாவது வரும் பந்தின் நிகழ்தகவு முதலில் வந்த பந்தைப் பொறுத்து அமைகிறது. முதலில் வந்த பந்து வெள்ளைப் பந்தாக வந்திருந்தால் இரண்டாவதாக வரும் பந்து பச்சையாக இருப்பதற்கான நிகழ்தகவு $\frac{5}{8}$ ஆக இருக்கும். முதலில் வந்த பந்து பச்சையாக இருந்திருந்தால் இரண்டாவது முறை பச்சைப் பந்து வருவதற்கான நிகழ்தகவு $\frac{4}{8}$ ஆக இருந்திருக்கும்.

இவ்வாறாக, முதல் நிகழ்வுக்கான நிகழ்தகவும், இரண்டாவது நிகழ்வின் நிகழ்தகவும் ஒன்றையொன்று சார்ந்திருக்கின்றன. எனவே இவை சார்புடைய நிகழ்ச்சிகள் ஆகும்.

சார்பிலா நிகழ்ச்சிகள் (Independent events)

பேதமையற்ற (unbiased) இரண்டு நாணயங்களைச் சுண்டிவிட்டால் ஒரு நாணயத்தில் வரக்கூடிய விளைவு மற்றொரு நாணயத்தில் வரக்கூடிய விளைவைப் பாதிப்பதில்லை. உதாரணத்திற்கு ஒரு ரூபாய் நாணயம் ஒன்றும், இரண்டு ரூபாய் நாணயம் ஒன்றும் சுண்டி விடப்பட்டால், ஒரு ரூபாய் நாணயத்தில் வரக்கூடிய நிகழ்ச்சி இரண்டு ரூபாய் நாணயத்தில் வரக்கூடிய நிகழ்ச்சியைப் பாதிக்காது. எனவே, அவை இரண்டும் சார்பிலா அல்லது தனித்த நிகழ்ச்சிகளாகும். இதுபோல பலவகையான சார்பிலா நிகழ்ச்சிகளை யோசித்துப் பார்க்கலாம்.

அதேபோல ஒரு நாணயத்தை இருமுறைகள் சுண்டி விட்டால், முதல்முறை நிகழ்ந்த நிகழ்ச்சி இரண்டாவது தடவை நாணயம் சுண்டப்பட்டபோது நிகழக்கூடிய நிகழ்ச்சியைப் பாதிக்காது. எனவே இவையும் சார்பிலா நிகழ்ச்சிகளே.

எளிய நிகழ்ச்சி (Simple event)

ஒரு சோதனையில் ஒரு நிகழ்ச்சி மட்டும் சாத்தியமானால் அது எளிய நிகழ்ச்சியாகும். ஒரு நாணயத்தைச் சுண்டி விட்டால், அதில் தலை அல்லது பூவில் ஏதேனும் ஒன்றுதான் நிகழும். எனவே, அது எளிய நிகழ்ச்சியாகும்.

கூட்டு நிகழ்ச்சிகள் (compound events)

ஒரே சமயம் இரண்டு அல்லது இரண்டுக்கு மேற்பட்ட நாணயங்களைச் சுண்டிவிட்டால், ஒவ்வொரு நாணயத்திலும் ஒரு நிகழ்வாக, ஒரே சமயம் பல மாதிரியான நிகழ்ச்சிகள் நிகழலாம். இரண்டு நாணயங்கள் சுண்டப்படும்போது

இரண்டும் பூவாகவோ, இரண்டும் தலையாகவோ, பூ ஒரு நாணயத்திலும், தலை இன்னொரு நாணயத்திலும் நிகழலாம். அப்படியிருப்பதால் அவை கூட்டு நிகழ்ச்சிகளாகும்.

நிகழ்தகவு கணக்கிடும் முறைகள்

மொத்த வாய்ப்புக்களில் சாதகமான வாய்ப்புக்கள் எந்த அளவு உள்ளதோ அதுவே நிகழ்தகவாகும். உதாரணத்திற்கு ஒரு நாணயம் சுண்டப்படும்போது அதில் தலை வருவதற்கான நிகழ்தகவு (½) அல்லது 0.5 ஆகும். இதில் நாணயங்களின் எண்ணிக்கை கூடக்கூட மொத்த வாய்ப்புக்களின் எண்ணிக்கை 2^n என்று கூடும். இரண்டு நாணயங்கள் என்றால், மொத்த வாய்ப்புக்கள் 2^2 என்று நான்கு ஆகும். அவை தத, பூபூ, தபூ, பூத என அமையும். பத்து நாணயங்கள் என்றால் 2^{10} ஆக மொத்த வாய்ப்புக்கள் இருக்கும். அவற்றையும் வரிசைப்படுத்த முயற்சித்துப் பார்த்தால், ஏன் அவற்றை இங்கு தரவில்லை என்ற கேள்விக்குப் பதில் கிடைக்கும்.

ஆனால், பல சமயங்களில் சாதகமான வாய்ப்பு எவ்வளவு, மொத்த வாய்ப்பு எவ்வளவு என்றறிவது இயலாததாகும். உதாரணத்திற்கு ஒரு மாணவர் ஒரு தேர்வில் தேர்ச்சியடைய என்ன நிகழ்தகவு என்று எப்படி அறிவது? ஒரு மாணவர் பல்கலைக்கழகத் தேர்வுகளில் தேர்ச்சி பெறுவதற்கு எந்தளவு வாய்ப்பு இருக்கிறதென்றால், நிச்சயம் தேர்ச்சி பெறுவேன் (P=1) என்பார்; அப்படியானால், போட்டித் தேர்வுகளிலும் நுழைவுத் தேர்வுகளிலும் தேர்ச்சி பெற எவ்வளவு வாய்ப்பு என்றால், நிச்சயம் இல்லை (P=0) என்பார். இம்மாதிரிச் சூழ்நிலைகளில் நிகழ்தகவு காண்பது கடினம். எனவே மனம்போன போக்கில் (subjective) ஏதேனும் ஒரு பதிலைச் சொல்வார்கள். இதை தன்னியல்பு நிகழ்தகவு (subjective probability) எனலாம்.

ஒரு நகரத்தில் ஓராண்டுக்கு எத்தனை பேர் விபத்துக்குள்ளாகிறார்கள் என்ற புள்ளி விபரம் இருந்தால், அதைப் பயன்படுத்தி, அந்த நகரத்தில் ஓராயிரம் பேரை சமவாய்ப்பு முறையில் எடுத்தால், அதில் எத்தனை பேர் விபத்துக்குள்ளாவார்கள் என்று ஓரளவு சரியாகச் சொல்ல முடியும் (இது எந்தவித புள்ளி விபரமும் இல்லாமல் தோராயமாகக் கூறுவதைவிடச் சரியாக இருக்கும்). இந்த மாதிரிப் புள்ளிவிபரங்களைக் கொண்டு கணிப்புகள் தொடர்ந்து நடத்தினால், அக்கணிப்புக்கள் கூடிய சீக்கிரம் உண்மையை ஒட்டி வரலாம்.

சேர்வைகள் (Combinations and Permutations)

மொத்த வாய்ப்புக்களைக் கணிக்க பல சமயங்களில் சேர்வைகள் பயன்படுத்தப்படுகின்றன. உதாரணத்திற்கு நான்கு மாணவர்களில் இருவரைத் தேர்ந்தெடுக்க வேண்டுமானால், மொத்த வாய்ப்புக்கள் நான்கு என்பது தவறாகும். நான்கு மாணவர்கள் அன்பு, ஆதி, இனியன், ஈகையன் (அ,ஆ,இ,ஈ) எனக் கொள்வோம். இவர்களில் இருவர் என்றால் அஆ, அஇ, அஈ, ஆஇ, ஆஈ மற்றும் இஈ என்று ஆறுவகையான கூறுகள் சேர்வைகளாக உள்ளன. எனவே, இந்த இரு நபர்கள் கொண்ட ஆறு கூறுகள்தான் மொத்த கூறுகளாகும். இதில் அன்பும் ஆதியும் வருவதற்கு நிகழ்தகவு = $\frac{1}{6}$ ஆகும், $\frac{2}{4}$ அல்ல. இதைப்போன்ற கூட்டு நிகழ்வுகளின்போது சேர்க்கைகள் மிகவும் பயன்படுகின்றன. இதில் இன்னொரு வகை வரிசைச் சேர்வை (permutation) ஆகும். இதில் வரிசை அல்லது இடத்திற்கும் முக்கியம் தரப்படுகிறது. மேலே கூறப்பட்ட உதாரணத்தில் அன்பும் ஆதியும் வரவேண்டும், அதிலும் ஆதி முதலிலும் அன்பு இரண்டாவதுமாக வரவேண்டும் என ஒரு நிபந்தனை கூடவே விதிக்கப்பட்டால், அப்பொழுது மொத்த வாய்ப்புக்கள் 12ஆகக் கூடிவிடும். அவை : அஆ, அஇ, அஈ,

ஆஇ, ஆஈ, இஈ, ஆஅ, இஅ, ஈஅ, இஆ, ஈஆ, ஈஇ ஆகும். இங்கு இடத்திற்கும் முக்கியத்துவம் அளிக்கப்படுகிறது. இதுவும் அவசியம். உதாரணத்திற்கு (1) (2) (3) (4) ஆகிய அட்டைகளை வைத்து 12 இரு இலக்க எண்கள் எழுதலாம். அவை : *12, 13, 14, 23, 24, 34, 21, 31, 41, 32, 42, 43.* இரண்டு எண்கள் இடம் மாறுவதால் அவற்றின் மதிப்புக்களும் மாறுகின்றன. இச்சூழ்நிலைகளில் வரிசைச்சேர்வை (permutation) பயன்படுத்தப்பட வேண்டும்.

சேர்வை மற்றும் வரிசைச் சேர்வையைக் காண கீழ்வரும் உதாரணங்கள் உதவலாம்.

$$n_{c_r} = \frac{n!}{(n-r)!\,r!}$$

$$n_{p_r} = \frac{n!}{(n-r)!}$$ (! ஃபேக்டோரியல் என அழைக்கப்படுகிறது)

5 மாணவர்களை இருவர் கொண்ட குழுக்களாகப் பிரித்தால் :

$$5C_2 = \frac{5!}{(5-2)!\,2!} = \frac{5\times4\times3\times2\times1}{3\times2\times1\times2\times1} = \frac{5\times4}{2\times1} = 10$$

5 மாணவர்களை மூவர் கொண்ட குழுக்களாகப் பிரித்தால் :

$$5C_3 = \frac{5!}{(5-3)!\,3!} = \frac{5\times4\times3\times2\times1}{2\times1\times3\times2\times1} = \frac{5\times4\times3}{1\times2\times3} = 10$$

10 மாணவர்களை இருவர் கொண்ட குழுக்களாகப் பிரித்தால்:

$$10C_2 = \frac{10\times9\times8\times7\times6\times5\times4\times3\times2\times1}{8\times7\times6\times5\times4\times3\times2\times1\times2\times1} = \frac{10\times9}{1\times2} = \boxed{45}$$

0 முதல் 4 வரையிலான 5 எண்களை வைத்து இரண்டு இலக்க எண்கள்

$$5P_2 = \frac{5!}{(5-2)!} = \frac{5 \times 4 \times 3 \times 2 \times 1}{3 \times 2 \times 1} = 20 \text{ எண்கள்}$$

1 முதல் 5 வரையிலான 5 எண்களை வைத்து மூன்று இலக்க எண்கள்

$$5P_3 = \frac{5!}{(5-3)!} = \frac{5 \times 4 \times 3 \times 2 \times 1}{2 \times 1} = 60 \text{ எண்கள்}$$

0 முதல் 9 வரையிலான 10 எண்களைக் கொண்டு இரண்டு இலக்க எண்கள்

$$10P_2 = \frac{10!}{(10-2)!} = \frac{10 \times 9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1}{8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1}$$

$$= 10 \times 9 = 90 \text{ எண்கள்}$$

இவ்வாறாகக் கூட்டு நிகழ்ச்சிகள் வரும்போது பொருத்தமான சேர்வைகளைப் பயன்படுத்தி மொத்த வாய்ப்பினைக் கணிக்கலாம்.

ஒரு பெட்டியில் 5 சிவப்புப் பந்துகளும் 3 வெள்ளைப் பந்துகளும் உள்ளன. அதிலிருந்து ஒரே சமயத்தில் 2 பந்துகள் எடுக்கப்படுகின்றன. அவையிரண்டும் சிவப்புப் பந்துகளாக இருப்பதற்கான நிகழ்தகவு யாது?

ஒரே சமயம் இரண்டு பந்துகள் எடுக்கப்படுவதால் இது ஒரு கூட்டு நிகழ்வு.

இதில் சாதகமான வாய்ப்புக்கள் $5c_2$. மொத்த வாய்ப்புக்கள் $8c_2$

$$5c_2 = \frac{5 \times 4}{1 \times 2} = 10 \;;\; 8c_2 = \frac{8 \times 7}{1 \times 2} = 28$$

எனவே நிகழ்தகவு $\frac{10}{28}$ ஆகும்.

கூட்டல் நியதி (Addition Theorem)

இரண்டு நிகழ்வுகள் சார்புள்ளதாகவோ, சார்பில்லாத தாகவோ, ஒன்றையொன்று விலக்கக்கூடியனதாகவோ இருக்கலாம். உதாரணத்திற்கு, ஒரு நாணயத்தைச் சுண்டும்போது தலையோ பூவோ வரலாம். அவற்றில் ஏதேனும் ஒன்றுமட்டும் தான் ஒரு சோதனையின்போது வரமுடியும். இப்படிப்பட்ட நிலையில் தலை அல்லது பூ வருவதற்கு நிகழ்தகவைக் கணக்கிட P(த) + P(பூ) எனக் கூட்ட வேண்டும். அது ½ + ½ = 1 என வரும். அதாவது, ஒருமுறை ஒரு நாணயத்தைச் சுண்டினால் பூவோ தலையோ நிச்சயம் வரும். இவ்விரு நிகழ்வுகளும் ஒன்றையொன்று விலக்குவதாக இருப்பதால், இரண்டும் நிகழ வாய்ப்பே இல்லை. எனவே இரண்டும் (பூவும் தலையும்) வருவதற்கு நிகழ்தகவு 0 (பூஜ்யம்). ஆகும்.

பெருக்கல் நியதி

வெவ்வேறு நாணயங்களில் பூவோ தலையோ வருவது சார்பிலா நிகழ்வுகளாகும். எனவே, இரண்டு நாணயங்கள் சுண்டப்படும்போது இரண்டிலும் தலை வருவதற்கான நிகழ்தகவு ½ × ½ = ¼ என்று ஆகிறது. இப்படிப்பட்ட சூழ்நிலையில் மொத்த வாய்ப்புக்கள்.

த, த;   த, பூ;   பூ, பூ;   பூ, த

இந்த நான்கு வாய்ப்புக்களில் இரண்டுமே தலையாக வர ஒரே ஒரு வாய்ப்புதான் உண்டு. எனவே இரண்டு நாணயங்களிலும் தலை வருவதற்கு ¼ நிகழ்தகவாகும்.

சார்பு மற்றும் சார்பிலா நிகழ்வுகளில் (A & B) ஏதேனும் ஒன்று வருவதற்கான நிகழ்தகவு காண : P(A) + P(B) - P(AB) ஒன்றையொன்று விலக்கும் நிகழ்வுகளானால் P(AB) = 0 என்பதனால் P(A) + P(B) என்று முடிகிறது.

உதாரணத்திற்கு 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 ஆகிய 10 எண்களில் ஓர் எண்ணை தேர்ந்தெடுத்தால் 2ஆல் மீதியின்றி வகுபடும் எண் வருவதற்கான நிகழ்தகவு = $\frac{5}{10}$; 5ஆல் மீதியின்றி வகுபடும் எண் வருவதற்கான நிகழ்தகவு $\frac{2}{10}$; 2 அல்லது 5ஆல் வகுபடும் எண் வருவதற்கான நிகழ்தகவு

$$= \frac{5}{10} + \frac{2}{10} - \frac{1}{10} = \frac{2+5-1}{10} = \frac{6}{10}$$

நிபந்தனை நிகழ்தகவு (Conditional Probability)

A, B என இரண்டு நிகழ்ச்சிகள் இருக்கின்றன. A முதலிலேயே நிகழ்ந்து விட்டதெனக் கொண்டு அப்படி இருந்தால், B நிகழ்வதற்கு என்ன வாய்ப்பு என்று கணிப்பது நிபந்தனை நிகழ்தகவு ஆகும். இதை p(B/A) அல்லது p(B given A) என்று அழைக்கலாம். Aயும் Bயும் சார்பிலா நிகழ்ச்சிகளாக இருந்தால் p(B/A) = p(B). Aயும் Bயும் சார்புடைய நிகழ்ச்சிகளாக இருந்தால் p (B/A) ≠ p(B). இதற்கு, ஒரு பெட்டியில் 4 வெள்ளைப் பந்துகளும் 5 பச்சை பந்துகளும் இருந்து, இரண்டாவதாக எடுக்கும் பந்து பச்சைப் பந்தாக இருப்பதற்கு என்ன நிகழ்தகவு என்று கேட்பதை எடுத்துக்காட்டாகக் கூறலாம். முதலில் எடுக்கும் பந்தை பெட்டியில் போட்டுவிட்டால் (With Replacement) முதல் மற்றும் இரண்டாவது நிகழ்ச்சிகள் சார்பிலா நிகழ்ச்சிகள் ஆகிவிடும். மாறாக, முதலில் எடுத்த பந்தை பெட்டிக்குள் போடாவிட்டால் (Without replacement) அது, இரண்டாவதாக எடுக்க இருக்கும் பந்தின் நிகழ்தகவைப் பாதிக்கும். உதாரணத்திற்கு முதலில் வந்த பந்து பச்சையென்றால் அதை மீண்டும் பெட்டிக்குள் போடாவிட்டால், இரண்டாவது பந்தை எடுக்கும்போது பெட்டிக்குள் மொத்தம் 8 பந்துகள் தான் இருக்கும் (முதலில் 9 பந்துகள் இருந்தன); பச்சைப் பந்துகள் 4 தான் (5 அல்ல) இருக்கும்.

முதலில் வந்த பந்து பச்சையாக இருந்திருந்தால் இரண்டாவது பச்சைப் பந்து வருவதற்கான நிகழ்தகவு $\frac{4}{8}$ ஆகும். இரண்டாவதாக வெள்ளைப்பந்து வருவதற்கான நிகழ்தகவு $\frac{4}{8}$ ஆகும். மாறாக, முதலில் வந்த பந்து வெள்ளைப் பந்தாக இருந்திருந்தால், இரண்டாவதாக பச்சைப்பந்து வருவதற்கான நிகழ்தகவு $\frac{5}{8}$ ஆகும்.

நிபந்தனை நிகழ்தகவினை விளக்க டேரோ யமனே (Taro Yamane, Statistics : An Introductory Analysis, Third Edition, Harper International Edition, New York, 1973, p. 113) கீழ்க்காணும் உதாரணத்தைப் பயன்படுத்துகிறார். ஒரு பானையில் 3 சிவப்புப் பந்துகளும், 7 பச்சைப் பந்துகளும் உள்ளன. 3 சிவப்புப் பந்துகளில் முதலாவது 1 என்றும் இரண்டாவது 2 என்றும் மூன்றாவது 3 என்றும் குறிக்கப்பட்டுள்ளன. அதுபோல 7 பச்சைப் பந்துகளில் 4 முதல் 10 வரை எண்கள் குறிக்கப்பட்டுள்ளன. இதில் ஒரு பந்து எடுக்கும்போது, அது சிவப்பாக இருப்பதற்கான நிகழ்தகவு என்ன என்று கேட்டால் $\frac{3}{10}$ என்றும் அது பச்சையாக இருப்பதற்கு நிகழ்தகவு என்ன என்று கேட்டால் $\frac{7}{10}$ என்றும் சொல்லலாம். ஒரு பந்து எடுக்கும்போது அந்தப் பந்தில் 5 குறிக்கப்பட்டுள்ளதற்கு என்ன நிகழ்தகவு என்றால் $\frac{1}{10}$ ஆகும். அதுபோல், 7 வருவதற்கு என்ன நிகழ்தகவு என்றால் அதுவும் $\frac{1}{10}$ ஆகும். ஆனால், மாறாக, எடுத்த பந்து பச்சையாக இருந்து அதில் 5 வருவதற்கான நிகழ்தகவு என்ன என்றால் $\frac{1}{7}$. அதாவது, பச்சைப்பந்துகள் 7 உள்ளன. அவற்றில்

ஒன்று 5 என்ற எண்ணைத் தாங்கியிருக்கும். எனவே $\frac{1}{7}$ ஆகும். இதை நிபந்தனைக்குட்படுத்தப்பட்ட நிகழ்தகவுப்படி

$$p(5 / green) = \frac{p(5 \& green)}{p(green)} = \frac{0.1}{0.7} = \frac{1}{7}$$

என்று டேரோ யமனே கூறுகிறார்.

அதேபோல, எடுத்த பந்து சிவப்பாக இருந்து அந்த பந்தில் 5 இருக்க நிகழ்தகவு எவ்வளவு? சிவப்புப் பந்தில் 1 முதல் 3 வரைதான் உள்ளன. எனவே, சிவப்புப் பந்தில் 5 வர வாய்ப்பே இல்லை. இதற்கான நிகழ்தகவு 0 (பூஜ்யம்) தான். இதை நிபந்தனை நிகழ்தகவு முறையைப் பயன்படுத்தி,

$$p(5 / red) = \frac{p(5 \& red)}{p(red)} = \frac{0.0}{0.3} = 0$$

என்கிறார் டேரோ யமனே.

அட்டவணை - 33

நிகழ்வுகள்	நிகழ்தகவு	கூட்டு நிகழ்வு	நிகழ்தகவு	நிபந்தனை நிகழ்தகவு
1	0.1			$0.1/0.3 = \frac{1}{3}$
2	0.1	சிவப்பு	0.3	
3	0.1			
4	0.1			
5	0.1			
6	0.1			
7	0.1	பச்சை	0.7	$0.1/0.7 = \frac{1}{7}$
8	0.1			
9	0.1			
10	0.1			
	1.0		1.0	

அட்டவணை - 34

நிகழ்வு	அலைவெண்	நிகழ்தகவு	கூட்டு நிகழ்வு	நிகழ்தகவு	நிபந்தனை நிகழ்தகவு
ஆண்நோயாளி	5	0.05	ஆண்	0.50	0.05/0.5 = 0.1
ஆண்நோயற்றவர்	45	0.45			0.45/0.5 = 0.9
					1.0
பெண்நோயாளி	10	0.10	பெண்	0.50	0.1/0.5 = 0.2
பெண்நோயற்றவர்	40	0.40			0.45/0.5 = 0.8
					1.0
மொத்தம்	100	1.00		1.0	1.0

அட்டவணை 34ம் நிபந்தனை நிகழ்தகவினை விளக்குவதற்காக சிறு வித்தியாசத்துடன் டேரோ யமனே கொடுத்துள்ளதே (பக்கம் 117) அட்டவணை 34ல் கொடுத்துள்ள விபரங்களிலிருந்து, தேர்ந்தெடுக்கப்பட்ட ஓர் ஆண், நோயாளியாக இருப்பதற்கான நிகழ்தகவு எவ்வளவு? என்பதைக் காணலாம். ஆண்கள் மொத்தம் 50 பேர்கள். அவர்களில் நோயாளிகள் 5 பேர்கள். எனவே நிகழ்தகவு $\frac{5}{50} = 0.1$. இதனை நிபந்தனை நிகழ்தகவாக

$$P\,(\text{நோயாளி} \,/\, \text{ஆண்}) = \frac{P\,(\text{நோயாளி} \,\&\, \text{ஆண்})}{P\,(\text{ஆண்})}$$

$$= \frac{0.05}{0.50} = 0.1 \quad \text{எனலாம்.}$$

நிபந்தனை நிக்ழ்வுக்கான சில சூத்திரங்கள்

$$P\,(B/A) = \frac{P\,(B \,\&\, A)}{P\,(A)} \quad \text{இதிலிருந்து,}$$

$$P(B \,\&\, A) = P(A)\,P(B/A)$$

$$P(A/B) = \frac{P\,(A\,\&\,B)}{P\,(B)}$$ இதிலிருந்து

$$P(A\&B) = P(B)\ \ P(A/B)$$

Aயும் Bயும் ஒன்றையொன்று விலக்கும் நிகழ்வுகளாக இருக்குமேயானால், P (B/A) = 0; P(A/B) = 0.

Aயும் Bயும் சார்பிலா நிகழ்வுகளாக இருக்குமேயானால், P(B/A) = P(B); P(A/B) = P(A).

அதாவது $$P(B/A) = \frac{P\,(B\,\&\,A)}{P\,(A)} = \frac{P(B)\,P(A)}{P\,(A)} = P(B)$$

$$P(A/B) = \frac{P\,(A\,\&\,B)}{P\,(B)} = \frac{P(A)\,P(B)}{P\,(B)} = P(A)$$

பேய்ஸ் கோட்பாடு (BAYES' THEOREM)

இதற்கு முன்னால் விவரிக்கப்பட்ட நிபந்தனை நிகழ்தகவுடன் நெருங்கிய தொடர்புள்ளது பேய்ஸ் கோட்பாடு ஆகும். பேனா செய்யும் இரண்டு இயந்திரங்கள் (H_1 and H_2) உள்ளன. இவற்றில் முதல் இயந்திரம் (H_1) 60 விழுக்காடு பேனாக்களைச் செய்கின்றது. மீதமுள்ள 40 விழுக்காடு பேனாக்களை இயந்திரம் இரண்டு (H_2) செய்கின்றது. மேலும், முதல் இயந்திரம் செய்கின்ற பேனாக்களில் 10 விழுக்காடு பழுதாகவும் (defective) இரண்டாவது இயந்திரம் செய்கின்ற பேனாக்களில் 20 விழுக்காடு பழுதாகவும் உள்ளன. பழுதாக உள்ளதற்கான வாய்ப்பினை B என்றும் பழுதாகாமல் இருப்பதற்கான வாய்ப்பினை A என்றும் குறிக்கலாம். இந்த விபரங்களிலிருந்து,

$$P(A/H_1) = 90\%;\ P(A/H_2) = 80\%;\ P(H_1) = 60\%;\ P(H_2) = 40\%$$

இந்த விபரங்களிலிருந்து பழுதல்லாத பேனா உற்பத்தி செய்யப்படுவதற்கான நிகழ்தகவு

$$P(A) = (0.9)(0.6) + (0.8)(0.4) = 0.86$$

ஒரு பழுதில்லாத பேனா முதல் இயந்திரத்தால் செய்யப்பட்டதற்கான நிகழ்தகவு என்ன என்பதை $P(H_1/A)$ என்று கேட்கலாம். அதாவது, ஒரு பழுதில்லாத பேனா கொடுக்கப்பட்டு இருந்து அது முதல் இயந்திரத்தால் செய்யப்பட்டதற்கான வாய்ப்பு என்ன என்று பொருள். இது முதல் இயந்திரத்தால் செய்யப்பட்ட பேனாவாக இருந்து பழுதல்லாத பேனாவாக இருக்க வாய்ப்பு என்ன $P(A/H_1)$ என்பதற்கு எதிர்மாறானதாகும். நிபந்தனை நிகழ்தகவு சூத்திரத்தின்படி,

$$P(H_1/A) = \frac{P(H_1 \& A)}{P(A)}$$

பேய்ஸ் (BAYES') கோட்பாட்டுப்படி

$$P(H_1/A) = \frac{P(A/H_1)\,P(H_1)}{\Sigma P(A/H_i)\,P(H_i)}$$

$P(H_1)$ = பேனா, முதல் இயந்திரத்தால் செய்யப்படுவதற்கான நிகழ்தகவு

$P(H_1/A)$ = பழுதற்ற பேனாவாக இருந்து அது முதல் இயந்திரத்தால் செய்யப்பட்டதற்கான நிகழ்தகவு

$P(H_1/A)$யையும் $P(H_2/A)$யையும் கீழ்வரும் அட்டவணை மூலம் கணக்கிடலாம்.

அட்டவணை - 35

நிகழ்வு	$P(H_i)$	$P(A/H_i)$	$P(H_i \& A)$	$P(H_i/A)$
இயந்திரம் 1(H_1)	0.6	0.9	0.54	0.54/0.86 = 0.63
இயந்திரம் 2(H_2)	0.4	0.8	0.32	0.32/0.86 = 0.37
மொத்தம்	1.0		0.86	1.00

$$P(H_1/A) = \frac{(0.9)(0.6)}{(0.54)+(0.32)} = \frac{P(A/H_1)P(H_1)}{P(H_1 \& A)+P(H_2 \& A)} = 0.63$$

$$P(H_2/A) = 0.37$$

நிகழ்தகவும் கணமும் (Probability and Set)

நிகழ்தகவுக் கோட்பாடுகளைப் புரிந்து கொள்ள கணங்களைப் பற்றிய கோட்பாடுகள் மிகவும் உதவுகின்றன. கணங்கள் பற்றிய கோட்பாடுகள் ஜார்ஜ் கேன்டர் (George Cantor : 1845-1918) என்பவரால் *1847*க்கும் *1895*க்கும் இடைப்பட்ட காலத்தில் விளக்கப்பட்டது. இவர் பல கணித வல்லுநர்களிடம் பயின்று 1867லேயே பெர்லின் பல்கலைக்கழகத்தில் (University of Berlin) முனைவர் பட்டம் (Ph.D) பெற்றவர்.

கணம் என்பது மிகச்சரியாக நிர்ணயித்து வரையறுக்கப்பட்ட பொருள்களின் / எழுத்துக்களின் / பெயர்களின், செயல்களின் ... தொகுப்பு (collection) ஆகும். அது மாணவர்கள் குழுவாகவோ, மரங்களாகவோ, மாடுகளாகவோ ... இருக்கலாம். ஒரு கணத்திற்குள் பல உட்கணங்கள் இருக்கலாம். எத்தனை உட்கணங்கள் (Subsets) இருப்பதென்பது அந்தக் கணத்தில் உள்ள கூறுகளின் எண்ணிக்கையைப் பொறுத்தது. எல்லாக் கணங்களுக்கும் ஒரு வெற்றுக்கணம் (null set) உட்கணமாக இருக்கும். வெற்றுக் கணம் ϕ எனக் குறிக்கப்படுகிறது.

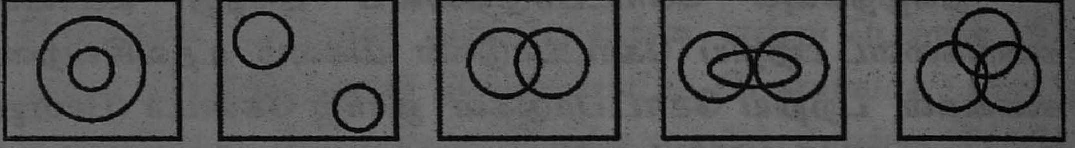
ஒவ்வொரு S_1ம் Sன் கூறு என்று இருந்தால், S_1 என்பது Sன் உட்கணம் எனப்படுகிறது.

(1, 2, 3) என்ற கணத்திற்கு

ϕ (1) (2) (3) (1, 2) (1, 3) (2, 3) (1, 2, 3) எனும் எட்டு வகையான உட்கணங்கள் இருக்கும். உட்கணங்களின் எண்ணிக்கையை 2^n என்று கணக்கிடலாம். இதில் n என்பது ஒரு கணத்தில் உள்ள கூறுகளின் எண்ணிக்கையாகும்.

கணங்களுக்கு இடையேயுள்ள தொடர்புகளை வென் (Venn) வரைபடம் (diagram) மூலமும் விளக்கலாம். இது யூலர் வரைபடம் (EULER DIAGRAM) என்றும் அழைக்கப்படுகிறது.

வரைபடம் - 29



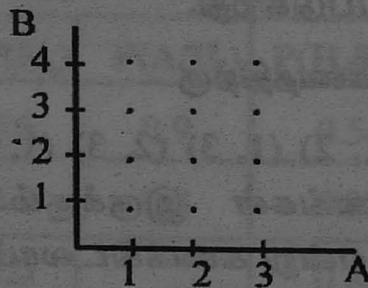
ஒரு கணத்தை மாதிரி வெளி (sample space) என்றும் சொல்லலாம். ஒரு சோதனையின்போது கிடைக்கும் நிகழ்வுகளை $E_1, E_2, E_3, ..., E_6$ என்றும் குறிக்கலாம். அல்லது $e_1, e_2, e_3, ..., e_6$ ஆகியவற்றை மாதிரிப் புள்ளிகள் (sample points) என்றும் குறிக்கலாம்.

மாதிரி வெளியை S என்று சொல்லி $S = \{e_1, e_2, e_3, e_4, e_5, e_6\}$ எனலாம்.

கார்டீசியன் பெருக்கல் (CARTESIAN PRODUCT)

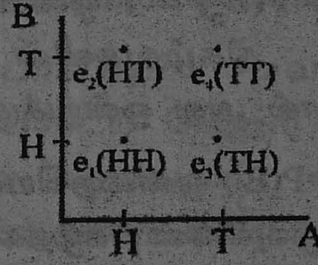
A என்ற மாதிரி 3 கூறுகளையும் (elements or sample points), B என்ற மாதிரி 4 கூறுகளையும் கொண்டிருந்தால், $A \times B$, 12 (3×4) வரிசைப்படுத்தப்பட்ட இணைகளைக் (ordered pairs) கொண்டிருக்கும். இதில் $A \times B$ கார்டீசியன் பெருக்கல் என அழைக்கப்படுகிறது. இதனை வரைபடம் 30இல் காணலாம்.

வரைபடம் - 30



எடுத்துக்காட்டாக, A,B இரண்டு நாணயங்களை ஒரே சமயம் சுண்டுவதை எடுத்துக் கொள்வோம். ஒவ்வொரு தடவை சுண்டும்போதும் இரண்டு விளைவுகள் உள்ளன. S_1 = (H,T), S_2 = HT. இதனைக் கார்டீசியன் பெருக்கல் மூலம் வரைபடம் 31இல் காட்டலாம். அங்கு நான்கு நிகழ்வுகள் (HH, HT, TT, TH) இருக்கும்.

வரைபடம் - 31



இதில் ஒவ்வொன்றும் மாதிரிப்புள்ளிகள். $S_1 \times S_2$ மாதிரிவெளி ஆகும்.

ஆறு சம பக்கங்களைக் கொண்ட இரண்டு பகடைக் கட்டைகளை உருட்டினால் அட்டவணை 35(அ)இல் கொடுத்துள்ளது போல மாதிரி வெளி (Sample Space) இருக்கும்.

அட்டவணை - 35 (அ)

		முதலாம் பகடைக்கட்டையில் தோன்றும் எண்கள்					
		1	2	3	4	5	6
இரண்டாம் பகடைக் கட்டையில் தோன்றும் எண்கள்	1	1,1	1,2	1,3	1,4	1,5	1,6
	2	2,1	2,2	2,3	2,4	2,5	2,6
	3	3,1	3,2	3,3	3,4	3,5	3,6
	4	4,1	4,2	4,3	4,4	4,5	4,6
	5	5,1	5,2	5,3	5,4	5,5	5,6
	6	6,1	6,2	6,3	6,4	6,5	6,6

நிகழ்தகவுக் கோட்பாட்டின் மாறாத உண்மைகள் (Axioms of Probability Theory)

1) $P(Ei) \geq 0$: ஒன்றை ஒன்று விலக்கக் கூடிய நிகழ்வுகள் இருக்கும் ஒரு சோதனையில், ஒரு நிகழ்வுக்கான நிகழ்தகவு எதிரிடை இல்லா (nonnegative) எண்ணாக இருக்கும்.

2) $P(E_1) + P(E_2) + \ldots + P(E_n) = 1$ ஒன்றை ஒன்று விலக்கும் எல்லா நிகழ்வுகளின் நிகழ்தகவையும் கூட்டினால் அது ஒன்றுக்குச் சமமாக இருக்கும்.

3) $P(E_i \text{ or } E_j) = P(E_i) + P(E_j)$ ஒன்றையொன்று விலக்கக்கூடிய இரு நிகழ்வுகளின் நிகழ்தகவுகளைக் கூட்டினால் அது ஒன்றுக்குச் சமமாக இருக்கும். இந்த மூன்று மாறாத உண்மைகளையும் அலைவெண் கோட்பாட்டின் அடிப்படையில் அட்டவணை 36இன் துணைகொண்டு விளக்கலாம்.

அட்டவணை - 36

நிகழ்வுகள்		அலைவெண்	சார்பு அலைவெண் (Relative frequency)
1	E_1	16	0.16
2	E_2	14	0.14
3	E_3	15	0.15
4	E_4	18	0.18
5	E_5	17	0.17
6	E_6	20	0.20
		100	1.00

1. $P(E_1) = 0.16$ .... $P(E_6) = 0.20$ அனைத்தும் எதிரிடை அல்லாத (non-negative) எண்கள்.

2. $P(E_1) + ... + P(E_6) = 0.16 + 0.14 + 0.15 + ... + 0.20 = 1.00$

3. $P(E_1 \text{ or } E_2) = P(E_1) + P(E_2) = 0.16 + 0.14 = 0.30$

கணித எதிர்பார்ப்பும் சமவாய்ப்பு மாறியும்
(Mathematical Expectation and Random variable)

கணித எதிர்பார்ப்பும் சமவாய்ப்பு மாறியும் நிகழ்தகவுடன் தொடர்புடைய மேலும் இரண்டு கருத்துக்கள் ஆகும். நிகழ்தகவில் கூறப்படும் எடையிட்ட கூட்டுச் சராசரியே ஒரு மாறியின் கணித எதிர்பார்ப்பு ஆகும்.

ஒரு பேதமற்ற நாணயத்தைச் சுண்டிவிடும்போது, அந்த நாணயத்தில் தலை மேலே தோன்ற நிகழ்தகவு ½யாகவும் பூ வருவதற்கு நிகழ்தகவு ½யாகவும் .இருந்து தலை வந்தால் ரூ.10 கிடைக்கும் என்றும் பூ வந்தால் ரூ.20 கிடைக்கும் என்றும் சொன்னால், ஒரு தடவை அந்த நாணயத்தை சுண்டுவதிலிருந்து எவ்வளவு பணம் எதிர்பார்க்கலாம்.

$$(10 \times 0.5) + (20 \times 0.5) = 5 + 10 = 15$$

எனவே, 15 ரூபாய் எதிர்பார்க்கலாம்.

மேலே கூறப்பட்ட விளையாட்டை விளையாட அனுமதிச்சீட்டு ரூபாய் 16க்குக் கிடைக்கும் என்றால் எத்தனை பேர் அந்த விளையாட்டுக்கு அனுமதிச் சீட்டு வாங்குவார்கள்? அப்படி யாரேனும் ஒருவர் அந்த விளையாட்டுக்கு அனுமதிச்சீட்டு வாங்கினால், அவர் எப்படிப்பட்டவராக இருப்பார்? அவர் நிச்சயம் துணிவோடு ஏற்றுக்கொள்பவராகத்தான் (RISK LOVER) இருப்பார். அவரைப் பொறுத்தமட்டில், அனுமதிச்சீட்டு வாங்குவதற்காக அவர் செலவு செய்யும் ரூ.16க்கான தியாகம் (sacrifice) அவர் அந்த விளையாட்டின் மூலம் கிடைக்கும் ரூ.15லிருந்து

பெறும் பயன்பாட்டை (utility) விடக் குறைவாக் இருக்க வேண்டும். இச்சூழ்நிலையில், நட்டத்தைத் தவிர்க்கும் ஒருவர் (risk averter) அந்த விளையாட்டுக்கு அனுமதிச்சீட்டு வாங்க முயற்சிக்க மாட்டார்.

முன்னர் கூறிய $(10 \times 0.5) + (20 \times 0.5) = 15$ என்பதை கணித எதிர்பார்ப்பாக $E(X) = x_1 p(x_1) + x_2\ p(x_2)$ என்று சொல்கிறார்கள்.

ஒரு சமவாய்ப்பு மாறி n அளவு விளைவுகளைக் கொண்டும் ஒவ்வொரு விளைவுக்கும் கிடைக்கும் வெகுமதி $x_1, x_2, ..., x_n$ என்றும் இருக்குமானால்

$$E(X) = x_1 p(x_1) + x_2 p(x_2) + ... + x_n p(x_n)$$ ஆகும்.

இக்கருத்துக்களைப் பயன்படுத்தி இப்பொழுது பல வியாபாரங்கள் நடத்தப்படுகின்றன. ஒரு குலுக்குச்சீட்டு வியாபாரத்தில் *1000* சீட்டுக்கள் விற்பனையாகியுள்ளன என்றும், ஒரு சீட்டின் விலை ரூ.*2* என்றும், அதில் வெற்றி பெற்றால் ரூ.*100* கிடைக்கும் என்றும், அந்த வியாபாரத்தில் வெற்றி பெற்ற ஒருவர்க்கு மட்டுமே ரூ.*100* கிடைக்கும் என்றும் கொண்டால்,

$$E(X) = (100 - 2)\ \frac{1}{1000} + (-2)\ \frac{999}{1000}$$

$$= \frac{98}{1000} - \frac{1998}{1000}$$

$$= 0.098 - 1.998 = -1.900$$

அதாவது சீட்டு வாங்குபவருக்கு *1* சீட்டுக்கு ரூ.*1.9* நட்டம், சீட்டு வியாபாரம் நடத்துபவருக்கு *1* சீட்டுக்கு ரூ.*1.9* இலாபம். எனவே பயன்பாட்டின் மூலம் மகிழ்ச்சி அடைபவர்தான் இச்சீட்டு வியாப்ரரத்தில் சீட்டு வாங்குவாரேயொழிய, பணம் வருமானத்தைக் கணக்கில் கொள்பவர் இந்தக் குலுக்கல் சீட்டை வாங்க மாட்டார் என்று சொல்லலாம்.

இந்தக் கருத்தைப் பயன்படுத்தி இயந்திரங்களுக்கோ அல்லது பொருள்களுக்கோ பராமரிப்புச் செலவு நிர்ணயிக்கப்படுகிறது. உதாரணத்திற்கு, ஓர் இயந்திரம் பழுதாவதற்கு என்ன வாய்ப்பு (எவ்வளவு நிகழ்தகவு) உள்ளது என்றும், அப்படிப் பழுதானால் அதைச் சரிசெய்ய எவ்வளவு செலவு ஆகும் என்றும் அந்த இயந்திரத்தை விற்பவருக்கு நன்றாகத் தெரியும்; அவருக்குள்ள அந்த இயந்திர அனுபவத்தால் இயந்திரம் விற்பவரின் அனுபவம் அளவுக்கு, அனுபவம் இல்லாமல் வாங்குபவருக்குச் சரியான விபரங்கள் தெரிய வாய்ப்பில்லை. இச்சூழ்நிலையில், ஓராண்டுக்குப் பராமரிப்புச் செலவை (Annual Maintenance Cost) நிர்ணயம் செய்யும்போது, (பொருள் விற்றதில் பெற்ற இலாபம் மட்டுமல்லாது) பொருளை விற்றவர் பராமரிப்புச் செலவை நிர்ணயம் செய்வதிலும் இலாபம் சம்பாதிக்க வாய்ப்பு உள்ளது.

உதாரணத்திற்கு ஓர் இயந்திரம், ஓராண்டில், பழுதாவதற்கு உள்ள வாய்ப்பும் (நிகழ்தகவும்) அப்படிப் பழுதானால் அதைச் சரிசெய்வதற்காகும் செலவும் கீழே கொடுக்கப்படுகிறது.

அட்டவணை - 37

கால இடைவெளி	முதல் 4 மாதங்கள்	இரண்டாம் 4 மாதங்கள்	மூன்றாம் 4 மாதங்கள்
பழுதாவதற்கான நிகழ்தகவு	¼	½	¼
செலவு	2000	1000	3000
எதிர்பார்ப்புச் செலவு	500	500	750

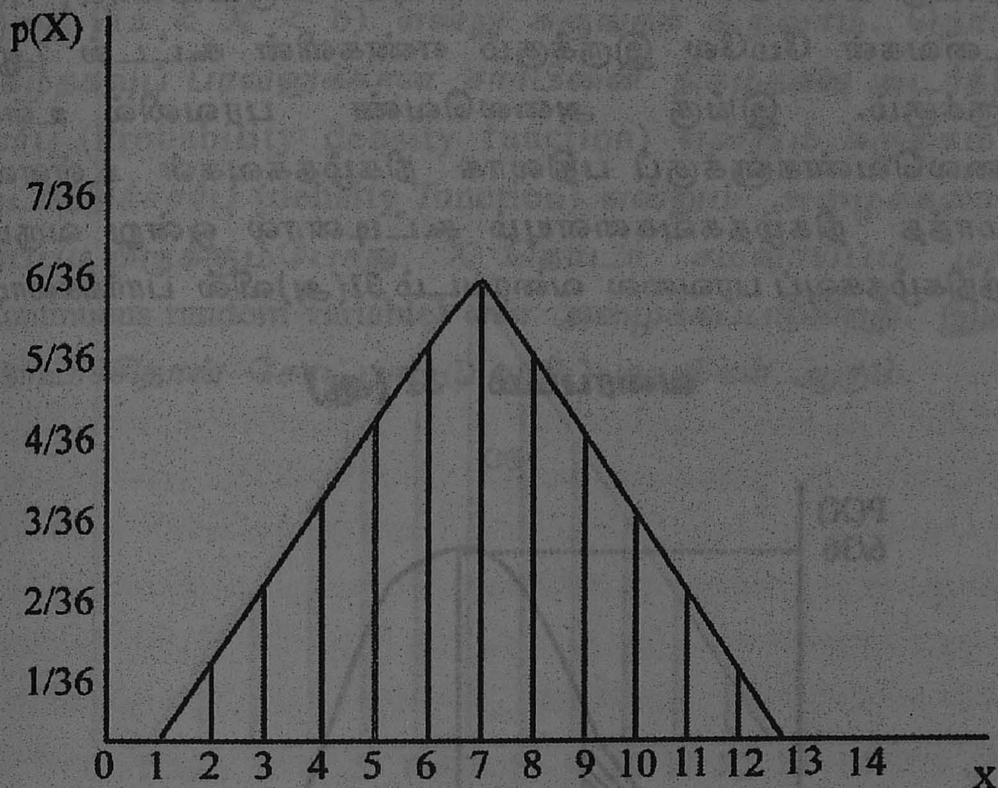
மொத்தச் செலவு 500 + 500 + 750 = 1750

இதற்கு, ஓராண்டு பராமரிப்புச் செலவு ரூ.1750க்கும் கீழ் நிர்ணயிக்கப்பட்டால், பொருளை வாங்கியவருக்கு இலாபம் வரவாய்ப்புள்ளது. ரூ.1750க்கும் மேல் ஓராண்டு பராமரிப்புச் செலவு நிர்ணயம் செய்யப்பட்டால் பொருளை விற்றவர் இலாபம் பெற வாய்ப்புள்ளது. பொருளை விற்றவர் இலாபம் பெற வாய்ப்புள்ளதை பொருளை வாங்கியவர் தெரிந்து கொண்டால், பொருளை வாங்கியவர் அறம் தவறும் ஆபத்தான (MORAL HAZARD) வழியில் செல்ல வாய்ப்புள்ளது. இப்படிப்பட்ட ஆபத்துக்கள் காப்பீடு (Insurance) வர்த்தகங்களில் நடக்க அதிக வாய்ப்புள்ளது. படிப்பறிவு மிக்கவர்கள் அதிகம் வாழும் நாடுகளில் ஒன்றான அமெரிக்காவில் 2008ஆம் ஆண்டு ஆகஸ்டு மாதம் தொடங்கிய பொருளாதார(நிதி) சீர்குலைவுக்கான தலையாய காரணங்களுள் ஒன்றாக அறம் தவறும் ஆபத்து இருந்ததாகக் கூறப்படுகிறது.

தனித்த நிகழ்தகவுப் பரவல்கள்
(Discrete Probability Distributions)

ஒரு குறிப்பிட்ட நிகழ்தகவுடன் பல முழு எண் கொண்ட மதிப்புக்களை எடுக்கும் ஒரு மாறியின் (variable) பரவலுக்கு தனித்த நிகழ்தகவுப் பரவல் என்று பெயர் (discrete probability distribution). ஓர் X என்ற மாறி p(X) என்ற சார்பைக் (function) கொண்டு ஒவ்வொரு Xக்கும் $(x_1, x_2, ..., x_n)$ நிகழ்தகவாக $p_1, p_2, ..., p_n$ என்று கொண்டிருந்தால் அதற்கு நிகழ்தகவுச் சார்பு (Probability function) என்றும் Xன் அலைவெண் சார்பு (frequency function) என்றும் பெயர். இந்த X, ஒரு குறிப்பிட்ட நிகழ்தகவுடன் பல மதிப்புக்களை எடுப்பதால் இது தனித்த சமவாய்ப்பு மாறி (discrete random variable) என்று அழைக்கப்படுகிறது. சமவாய்ப்பு மாறி சந்தர்ப்ப மாறியென்றும் (Chance variable) எதிர்பாரா மாறியென்றும் (stochastic variable) அழைக்கப்படுகிறது. உதாரணத்திற்கு, அட்டவணை 38ஐப் பார்க்கலாம்.

வரைபடம் - 31 (அ)



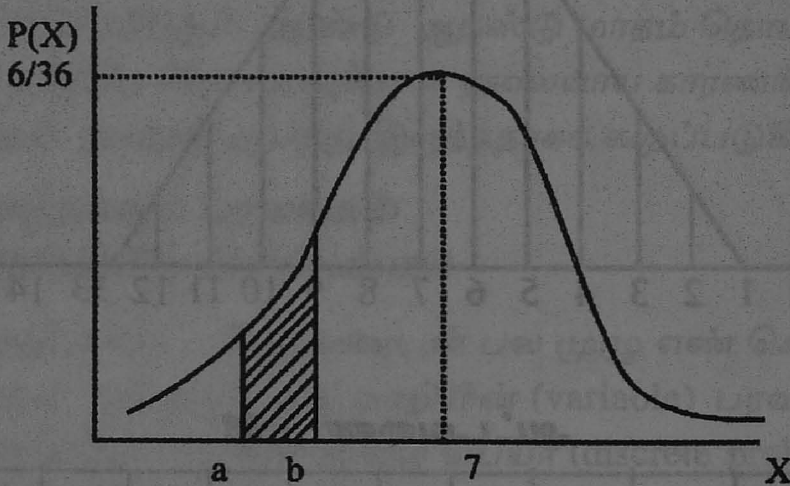
அட்டவணை - 38

X	2	3	4	5	6	7	8	9	10	11	12
P(X)	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

அட்டவணை 38இல் கொடுக்கப்பட்டுள்ளவை இரண்டு பேதமற்ற (fair) பகடைக் கட்டைகளை (ஆறு சம பக்கங்கள் கொண்டவை) உருட்டும்போது எதிர்பார்க்கக்கூடிய நிகழ்வுகளும் அவை நிகழ்வதற்கான நிகழ்தகவுகளும். உதாரணத்திற்கு, இரண்டு பகடைக்கட்டைகளின் (dice) மேலும் தோன்றும் எண்களைக் கூட்டினால் 7ஆகக் கிடைப்பதற்கான

நிகழ்தகவு $\frac{6}{36} = \frac{1}{6}$. எனவே, 600 தடவைகள் அந்த இரண்டு பகடைக் கட்டைகளையும் உருட்டினால், 100 தடவைகள் மேலே இருக்கும் எண்களின் கூட்டல் 7ஆக இருக்கும். இங்கு அலைவெண் பரவலில் உள்ள அலைவெண்களுக்குப் பதிலாக நிகழ்தகவுகள் உள்ளன. மொத்த நிகழ்தகவுகளையும் கூட்டினால் ஒன்று வரும். இந்நிகழ்தகவுப் பரவலை வரைபடம் 31(அ)வில் பார்க்கலாம்.

வரைபடம் - 31 (ஆ)



வரைபடம் 31 (அ)வைப் போலவே தொடர் நிகழ்தகவுப் பரவலுக்கும் (Continuous probability distribution) ஒரு வரைபடம் வரையலாம். அதுதான் வரைபடம் 31 (ஆ).

தொடர்நிகழ்தகவுப் பரவலில் X எனும் மாறி முழு எண்களை மட்டுமல்லாது பின்ன எண்களையும், தசம எண்களையும் ஏற்கும். வரைபடம் 31 'அ' விலும் 31'ஆ' விலும் வளைகோட்டிற்குள் இருக்கும் மொத்தப் பரப்பு மொத்த நிகழ்தகவாகும், அதாவது 1. வரைபடம் 31'ஆ' வில்

நிழலாக்கப்பட்ட பகுதி Xன் மதிப்பு a முதல் b வரையிருப்பதற்கான நிகழ்தகவினைக் காட்டுகிறது. இதை $P(a < X < b)$ என்று சுருங்கக் கூறலாம். தொடர் நிகழ்தகவுப் பரவலுக்கான சார்பினை நிகழ்தகவு அடர்த்திச் சார்பு (Probability density function) என்றும் சுருக்கமாக அடர்த்திச்சார்பு (density function) என்றும் அழைக்கலாம். இப்படியிருக்கும்போது X தொடர் சமவாய்ப்பு மாறி (Continuous random variable) என அழைக்கப்படுகிறது. இதன் தொடர்ச்சிதான் கோட்பாட்டு வழிப்பரவல்கள் ஆகும்.

ശ

8. கோட்பாட்டுவழிப் பரவல்கள்
(THEORETICAL DISTRIBUTIONS)

இதற்கு முன்னர் பல இடங்களில் விவரித்துள்ளதுபோல் மாறிகள் பலவகைகள் உள்ளன. அவற்றில் இரண்டுவகை மாறிகள் இங்கு பேசப்படுகின்றன. ஒன்று, தனித்த சமவாய்ப்பு மாறிகள் (discrete random variables). மற்றொன்று, தொடர் சமவாய்ப்பு மாறிகள் (Continuous random variables) தனித்த சமவாய்ப்பு மாறிகளுக்கு உதாரணமாக நாணயத்தைச் சுண்டும்போது தலை வருவதைச் சொல்லலாம். அல்லது ஒரு நாளைக்கு ஒரு நகரத்தில் நடக்கும் விபத்துக்களைக் கூறலாம். அல்லது ஒரு நாளைக்கு எத்தனை பேருக்கு பாம்பு கடிக்கிறது என்று சொல்லலாம். இதுபோல ஒரு தேர்வில் தேர்ச்சி பெறும் மாணவர்களையும் தனித்த சமவாய்ப்பு மாறிகளுக்கு எடுத்துக்காட்டாகக் கூறலாம். இவையனைத்தும் (தலை, விபத்துக்கள், மனிதர்கள், மாணவர்கள்) முழு எண்களாகத்தான் வரும். அரை என்றோ கால் என்றோ (0.5 என்றோ 0.25 என்றோ) வராது. ஆனால் இவைகளுக்குள் இந்த ஒற்றுமை இருந்தாலும் ஒரு வேற்றுமையும் உண்டு. அது அந்த நிகழ்வுகளில் சில அடிக்கடி நிகழலாம்; சில எப்போதாவது நிகழலாம். சில சோதனைகளில் இரண்டே நிகழ்வுகளுக்கு வாய்ப்பு இருக்கலாம். உதாரணத்திற்கு ஒரு நாணயத்தைச் சுண்டும்போது பூ அல்லது தலையாகிய இரண்டு வாய்ப்புகள்தான் உள்ளன. சில நிகழ்வுகளுக்கு சமவாய்ப்பு இருக்கலாம். உதாரணத்திற்கு ஒரு நாணயத்தைச் சுண்டும்போது தலை வருவதற்கும் பூ வருவதற்கும் சம வாய்ப்புக்கள் இருக்கலாம். ஆனால் ஒரு மாணவர் ஒரு பல்கலைக்கழகத் தேர்வில் தேர்ச்சி அடைய மிக அதிக வாய்ப்பும், தோல்வியடைய மிகக் குறைவான வாய்ப்பும்

இருக்கலாம். ஆனால், அதே சமயம் ஒரு மாணவர் போட்டித் தேர்வில் தேர்ச்சி பெற்று ஒரு நல்ல பணியில் அமர வாய்ப்பு மிகக் குறைவாகவும் அப்படிப்பட்ட தேர்வில் தோல்வியுற அதிக வாய்ப்பும் இருக்கலாம். எனவே இப்படிப்பட்ட வித்தியாசமான சூழ்நிலைகளில் அவற்றிற்கான அலைவெண் பரவல்களும், நிகழ்தகவுப் பரவல்களும் மாறுபடும். இச்சூழ்நிலைகளை விளக்குவதற்கு வெவ்வேறு கோட்பாடுகள் உள்ளன.

அதுபோல சில மாறிகள் முழு எண்கள் மட்டுமல்லாது எல்லாவிதமான (பின்னம், தசம எண்கள்) எண்களையும் ஏற்கும். உதாரணத்திற்கு, குழந்தைகளின் எடை 2.0023 கிலோவாகவோ 2½ கிலோவாகவோ இருக்கலாம். ஒரு மாணவர் பெறும் மதிப்பெண் 55.37 ஆகவோ $6\frac{1}{3}$ ஆகவோ இருக்கலாம். இவ்வாறு மதிப்புக்கள் இருக்கும்போது, அவற்றின் நிகழ்தகவுப் பரவல்கள் (Probability distributions) அல்லது அடர்த்திப் பரவல்கள் (Density distributions) வெவ்வேறாக இருக்கலாம்.

எனவே, இவைகளைப் பற்றிப் புரிந்து கொள்ள வெவ்வேறு கோட்பாடுகள் உள்ளன. அவற்றில் மூன்றினை மட்டும் இங்கு காணலாம்.

1. ஈருறுப்புப் பரவல் (Binomial distribution)

2. பாய்ஸான் பரவல் (Poisson distribution)

3. இயல்நிலைப் பரவல் (Normal distribution)

பல்லுறுப்புப் பரவல் (Multinomial distribution), அதிபெருக்குப் பரவல் (Hypergeometric distribution), χ^2 பரவல் (χ^2 distribution), t பரவல் (t distribution) மற்றும் F பரவல் (F distribution) ஆகியவையும் உள்ளன. மேலே கூறப்பட்டுள்ள முதல் மூன்று பரவல்களுள் எப்படிப்பட்ட சூழ்நிலை எந்தப் பரவலுக்குள்

வரும் என்பதையும் எந்தப் பரவல் எப்படிப்பட்ட சூழ்நிலையைக் கொண்டிருக்கும் என்பதையும் அறியலாம்.

கீழ்வரும் மூன்று நிலைகளும் ஒருங்கே இணைந்து இருந்தால் அதற்கு ஈருறுப்புப் பரவலைப் பயன்படுத்தலாம்.

1. ஒவ்வொரு முயற்சியிலும் (trial) இரண்டு நிகழ்வுகள் (events) மட்டுமே இருக்க வேண்டும்.
2. இரண்டு நிகழ்வுகள் நிகழ்வதற்கும் சமமான வாய்ப்பு இருக்க வேண்டும்.
3. ஒரு குறிப்பிட்ட நிகழ்தகவுடன் நிகழக்கூடிய மாறியின் மதிப்புக்கள் முழு எண்களாக இருக்க வேண்டும்.

உதாரணத்திற்கு, ஒரு நாணயத்தைச் சுண்டிவிடும் சோதனையில் இரண்டு நிகழ்ச்சிகள் மட்டுமே நிகழும். அவை தலை மற்றும் பூ. இது முதலாவது நிலை. இரண்டாவதாக, நல்ல நாணயத்தில் உள்ள இரண்டு நிகழ்ச்சிகளுள் ஒவ்வொன்றும் நிகழ்வதற்கு (தலை மற்றும் பூ) சமமான (equally likely) வாய்ப்புக்களே உள்ளன. இது இரண்டாவது நிலை. நிகழக்கூடிய நிகழ்ச்சிகள் தலை என்றாலும் பூ என்றாலும் ஒன்றாகத்தான் இருக்க முடியுமேயொழிய *0.75* தலையோ *1.25* தலையோ நிகழ வாய்ப்பில்லை. இது மூன்றாம் நிலை.

பகடைக்கட்டையை வைத்து இதற்கு உதாரணம் சொல்ல வேண்டுமானால், பகடைக் கட்டையில் ஆறு வெவ்வேறு விதமான நிகழ்ச்சிகளுக்கு *(1, 2, 3, 4, 5, 6)* வாய்ப்பு இருந்தாலும், ஒற்றைப் படை எண் *(1, 3, 5)* இரட்டைப்படை எண் என்று கொண்டால் அங்கு இரண்டு வகையான நிகழ்வுகள்தான் உள்ளன. [ஆறுவகையான நிகழ்ச்சிகளையும் அப்படியே கொண்டால் அதை பல்லுறுப்புப் பரவல் (Multinomial distribution) என்றழைக்கலாம்]. இது முதல் நிலை. ஒற்றைப்படை எண் நிகழ்வதற்கும் இரட்டைப்படை

எண் நிகழ்வதற்கும் சமமான வாய்ப்பே உள்ளது ($\frac{3}{6}$, $\frac{3}{6}$). இது இரண்டாம்நிலை. மாறியின் மதிப்புக்கள் ஒற்றைப்படை எண்களாகிய 1, 3, 5 ஆகவோ, இரட்டைப்படை எண்களாகிய 2, 4, 6 ஆகவோ இருக்குமேயன்றி 1½ என்றோ 2.5 என்றோ வர வாய்ப்பில்லை. இது மூன்றாம் நிலை.

அன்றாட வாழ்விலிருந்தும் இதற்கான உதாரணங்கள் எடுக்கலாம். ஒரு பணியிடத்தை நேர்காணல் மூலம் நிரப்புவதற்கு ஒரு நிறுவனம் செய்கின்ற செயலை எடுத்துக் கொள்ளலாம். அன்பு, ஆதி என்ற சம திறமை கொண்ட இருவர் நேர்காணலுக்குச் செல்கிறார்கள். அங்கு இரண்டு வாய்ப்புக்களே உள்ளன; ஒன்று தெரிவு செய்யப்படலாம் அல்லது தெரிவு செய்யப்படாமல் போகலாம். எனவே, இரண்டு நிகழ்வுகள் நடப்பதற்குத்தான் வாய்ப்புள்ளது. இது முதல் நிலை. இருவரும் சமதிறமை கொண்டவர்களாக உள்ளார்கள் என்பதால் இருவருக்குமே சம வாய்ப்பு உள்ளது. அன்புக்கு இடம் கிடைப்பதற்கும், கிடைக்காததற்கும் சமவாய்ப்பு உள்ளது. அது போல ஆதிக்கும். இது இரண்டாம் நிலை. இதில் அன்பு என்ற ஒருவருக்கோ ஆதி என்ற ஒருவருக்கோ பணியிடம் கிடைக்கும் என்று சொல்லலாமேயொழிய, 1½ மனிதர் என்றோ 1.75 மனிதர் என்றோ சொல்ல முடியாது. இது மூன்றாம் நிலை. இதில் இருவருமே சமவாய்ப்பு உள்ளவர்கள், சமதிறமை உள்ளவர்கள் என்று அநுமானம் செய்துள்ளோம். அன்புக்கு இடம் கிடைப்பதற்கும் கிடைக்காததற்கும் சம வாய்ப்பு உள்ளது என்கிறோம். இவை இயல்பு வாழ்வில் சாத்தியமா என்பது கேள்விக்குரியதுதான். ஆனால் ஒரு நாணயத்தில் தலையும் பூவும் சமவாய்ப்பு நிகழ்ச்சிகள் என்றால் ஒத்துக் கொள்ளலாம்.

மேலே கூறிய உதாரணங்களில், நிகழக்கூடிய இரு நிகழ்வுகளும் சம வாய்ப்பு நிகழ்வுகள் என்று எடுத்துள்ளோம்.

ஆனாலும், முன்னரே கூறியபடி சமவாய்ப்பு இல்லாத நிகழ்வுகளும் இருக்கின்றன. முன்னர் கூறிய மூன்று நிலைகளில், முதல் மற்றும் மூன்றாம் நிலைகள் சரியாக இருந்து, இரண்டாம் நிலையில் கூறிய சமவாய்ப்பு இல்லாமல், இரு நிகழ்வுகளும் நிகழக்கூடிய வாய்ப்புக்களில் மிகவும் வித்தியாசம் இருந்தால் அப்படிப்பட்ட பரவலை பாய்ஸான் பரவல் (Poisson Distribution) என்று சொல்லலாம்.

முன்னர் கூறப்பட்ட மூன்று நிலைகளில் மூன்றாம் நிலையான முழு எண்கள் இல்லாமல், மாறிகளின் மதிப்புக்கள் தொடர் எண்களில் (continuous) எதை வேண்டுமானாலும் ஏற்கும் என்றிருந்தால் அதை இயல்நிலைப் பரவல் (Normal distribution) எனலாம்.

மேலே விளக்கப்பட்ட மூன்று வகையான பரவல்களில், இதுவரையிலும் மிக அதிகமாகப் பயன்படுத்தப்பட்ட உதாரணமாகிய நாணயத்தைச் சுண்டுதலை எடுத்து விளக்கக் கூடிய ஈருறுப்புப் பரவலை முதலில் எடுத்துக் கொள்வோம்.

ஈருறுப்புப் பரவல் (Binomial Distribution)

கோட்பாட்டு வழிப் பரவல்களில் (Theoretical distributions) அல்லது எதிர்பார்க்கப்பட்ட அலைவெண்களையுடைய பரவல்களில் (Expected Frequency Distributions) முதலாவதாக இங்கு ஈருறுப்புப் பரவலைக் காணலாம். இப்பரவலை முதன்முதலில் தருவித்த கணித மேதையான ஜேக்கப் பெர்னோலி (JAKOB BERNOULLI : 1654-1705) அவர்களின் பெயருடன் பெர்னோலி பரவல் என்றும் அழைக்கிறார்கள்.

இரண்டு நிகழ்ச்சிகளாக ஒற்றைப்படை எண், இரட்டைப்படை எண் என்றோ; தலை, பூ என்றோ; வெற்றி, தோல்வி என்றோ; ஆண், பெண், என்றோ; இடது, வலது என்றோ; மேல், கீழ் என்றோ கொள்ளலாம். வெற்றி, தோல்வி எனும் இரண்டு நிகழ்வுகளை இங்கு உதாரணமாக

எடுத்துக் கொள்வோம். பொதுவாக வெற்றியை P என்றும் தோல்வியை Q என்றும் சுருக்கமாக எடுத்துக் கொள்வோம். எனவே 1-P = Q; P + Q = 1; 1 - Q = P. அதுபோல, ஒரு குழந்தை பிறப்பில் பெண் குழந்தை பிறப்பதற்கான நிகழ்தகவை (G) ½ என்று கொண்டால் ஆண்குழந்தை பிறப்பதற்கான நிகழ்தகவை (B) ½ என்று கொள்ளலாம். பொதுவாக ஆண் அல்லது பெண் குழந்தை பெறுவதற்கு வாய்ப்பு அதிகம். இங்கு இருபாலிலும் சேராத குழந்தை பிறக்க வாய்ப்பு மிகக்குறைவாகவே உள்ளதால், வாய்ப்பு இல்லையென அநுமானிக்கிறோம். இவ்வாறாக ஒரு நிகழ்வு நிகழ்வதற்கான நிகழ்தகவை p(X) என்று கொண்டால்,

$$p(X) = {}^{n}c_{X}p^{X}q^{n-X} = \frac{n!}{(n-r)!\,x!}\,p^{X}q^{n-X}$$

X, 1 முதல் n வரை எதுவானாலும் இருக்கலாம்.

$$n! = n\,(n-1)\,(n-2)\,\ldots\,(n - n+1)$$

0! = 1 என வரையறுக்கப்பட்டுள்ளது.

எடுத்துக்காட்டாக, நான்கு குழந்தைகள் உள்ள *1600* குடும்பங்களில் எத்தனை குடும்பங்களில் சரியாக 2 பெண் குழந்தைகள் 2 ஆண் குழந்தைகள் இருப்பார்கள்? பெண்குழந்தை பெற வாய்ப்பு ½ என்றும் ஆண்குழந்தை பெற வாய்ப்பு ½ என்றும் கொள்வோம்.

இது ஈருறுப்புப் பரவலை ஒத்தது. ஏனெனில் இரண்டு வாய்ப்புக்கள் உள்ளன. இரண்டுமே சமவாய்ப்புக்கள்; குழந்தைகள் தசம எண்களாகவோ பின்னங்களாகவோ குறிப்பிடப்படுவதில்லை. முதலில் 2 பெண் குழந்தைகள் இருக்க நிகழ்தகவு என்ன என்று கணிக்கலாம்.

$$p(X) = {}^{n}c_{X}p^{X}q^{n-X} = P\,(2\text{ பெண்}) = 4c_{2}p^{2}q^{2}$$

$$= 6\left(\frac{1}{2}\right)^{2}\left(\frac{1}{2}\right)^{2} = \frac{6}{16}$$

எனவே 4 குழந்தைகள் உள்ள ஒரு குடும்பத்தில் சரியாக இரண்டு பெண் குழந்தைகள் இருப்பதற்கான நிகழ்தகவு $\frac{6}{16}$ ஆகும். எனவே, 1600 வீடுகளில், 600 வீடுகளில் $\left(\frac{6}{16} \times 1600\right)$ சரியாக 2 பெண்குழந்தைகள் இருப்பார்கள் என்று சொல்லலாம். வீடுகள் சரியான முறையில் சமவாய்ப்பு மாதிரி மூலம் தேர்வு செய்யப்பட்டிருந்தால் இது சரியாக இருக்க வாய்ப்பு அதிகம். இதுபோல பெண்குழந்தையில்லாத வீடுகள் (0 பெண்கள்), 1 பெண் குழந்தையுள்ள வீடுகள், 3 பெண் குழந்தைகள் உள்ள வீடுகள், 4 பெண்குழந்தைகள் உள்ள வீடுகள் எத்தனை எனவும் கணக்கிடலாம். இந்தக் கணிப்பு அனேகமாகச் சரியாக இருக்கும். இவ்வாறாக, ஈருறுப்புப் பரவல் பயன்படுகிறது. எனவே, பலவிதமான கொள்கைகள் வகுக்கத் தேவைப்படும் புள்ளி விபரங்களை ஈருறுப்புப் பரவல் முறையைப் பயன்படுத்திப் பெறலாம்.

இன்னொரு உதாரணமும் எடுக்கலாம். நான்கு பேதமற்ற நாணயங்களை 160 முறைகள் சுண்டினால், எத்தனை தடவைகள் அங்கே 3 தலைகள் அல்லது 3 தலைகளுக்கு மேல் வர வாய்ப்புள்ளது?

இதற்கும் p(3 தலைகள்) + p(4 தலைகள்) எனக் கொண்டு விடையைக் கண்டுபிடிக்கலாம்.

$$p(3\ \text{தலைகள்}) = 4c_3p^3q^1 = \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^1 = \frac{4}{16}$$

$$p(4\ \text{தலைகள்}) = 4c_4p^4q^0 = \left(\frac{1}{2}\right)^4 \left(\frac{1}{2}\right)^0 = \frac{1}{16}$$

இதிலிருந்து 40 தடவைகள் $\left(\frac{4}{10} \times 160\right)$ 3 தலைகளும் 10 தடவைகள் $\left(\frac{1}{16} \times 160\right)$ 4 தலைகளும் வர வாய்ப்பு இருக்கின்றது என்று கூறலாம். உண்மையிலேயே ஒரு

சோதனை செய்து பார்த்தால் இங்கே கூறப்பட்ட பதில்கள் எவ்வளவுக்குச் சரியாக உள்ளதெனக் காணலாம்.

தலை வருவது ஒரு சமவாய்ப்பு மாறி என்று கொண்டு 3 பேதமற்ற நாணயங்களைச் சுண்டினால் கிடைக்கும் பரவலின் கூட்டுச் சராசரியை எப்படிக் கணிப்பது என்று இனி காணலாம். இதில் 'தலை' வரும் நிகழ்தகவும் 'பூ' வரும் நிகழ்தகவும் சமம் (அதாவது ½, ½) என்று கொள்வோம். அப்படிக் கொண்டால் அட்டவணை 39 கிடைக்கும்.

அட்டவணை - 39

தலைகளின் எண்ணிக்கை	p(X)	(X) p(X)
0	$^nc_0p^0q^{n-0} = \frac{1}{8}$	0
1	$^nc_1p^1q^{n-1} = \frac{3}{8}$	$\frac{3}{8}$
2	$^nc_2p^2q^{n-2} = \frac{3}{8}$	$\frac{6}{8}$
3	$^nc_3p^3q^{n-3} = \frac{1}{8}$	$\frac{3}{8}$

$$E(X) = \Sigma(X)\, p(X) = \frac{12}{8} = 1.5$$

1.5 தான் 3 பேதமற்ற நாணயங்களைச் சுண்டும்போது கிடைக்கும் பரவலின் கூட்டுச் சராசரியாகும். இது $3 \times \frac{1}{2}$ ஆகும். அதாவது 3 நாணயங்கள், தலை வருவதற்கு நிகழ்தகவு $\frac{1}{2}$. அதுபோல, நான்கு பேதமற்ற நாணயங்களைச் சுண்டினால் கிடைக்கும் பரவலின் கூட்டுச்சராசரி 2 ஆக ($4 \times \frac{1}{2}$) இருக்கும். எனவே, ஓர் ஈருறுப்புப் பரவலின் கூட்டுச் சராசரியை np என்று கொள்ளலாம். இது தொடர்பான மற்ற பண்பளவைகள் அட்டவணை 40இல் தரப்படுகின்றன.

அட்டவணை - 40

ஈருறுப்புப் பரவலின் குணாதிசயங்கள்

கூட்டுச் சராசரி	(Mean)	$\mu = np$
மாறுபாடு	(Variance)	$\sigma^2 = npq$
திட்டவிலக்கம்	(Standard deviation)	$\sigma = \sqrt{npq}$
கோட்டத்தின் அசைவுக்கெழு	(Moment coefficient of Skewness)	$\alpha_3 = \frac{q-p}{\sqrt{npq}}$
தட்டைத் தன்மையின் அசைவுக்கெழு	(Moment coefficient of Kurtosis)	$\alpha_4 = 3 + \frac{1-6pq}{npq}$

ஈருறுப்புப் பரவலைக் காணல் (Fitting binomial distribution)

அட்டவணை 41இல் கொடுக்கப்பட்டுள்ள புள்ளி விபரங்களுக்கு எப்படி ஈருறுப்புப் பரவலைக் காண்பது என்று இனி தெரிந்து கொள்ளலாம். இதற்கு எடுத்துக் கொள்ளப்பட்ட முயற்சி 5 நாணயங்களைச் சுண்டுவதாகும்.

அட்டவணை - 41

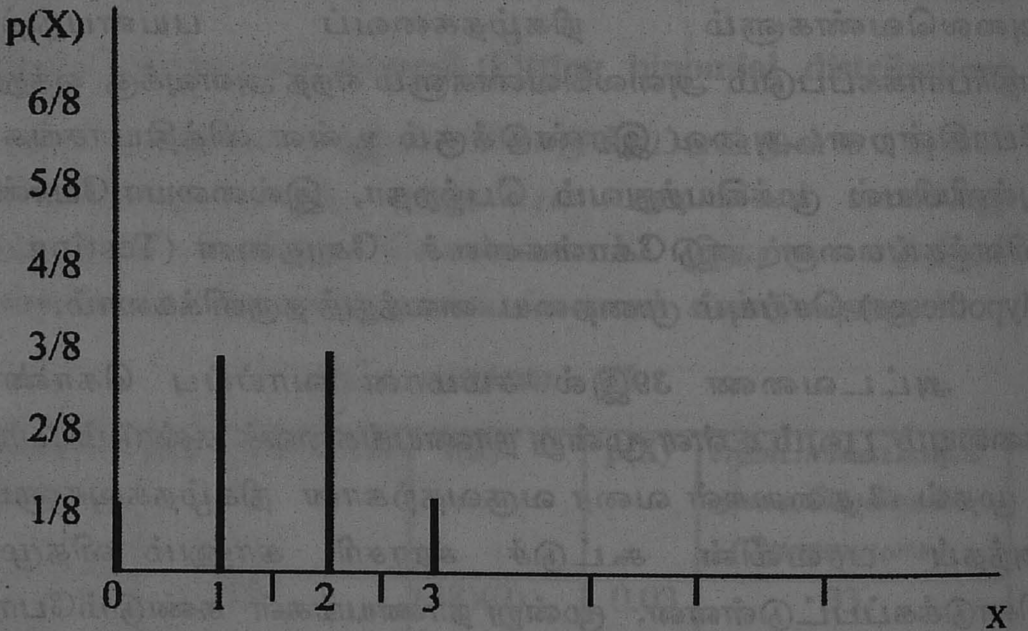
தலைகளின் எண்ணிக்கை	அலைவெண்கள் (f)	f(X)	p(X)	எதிர்பார்க்கப்படும் அலைவெண்கள் (தோராயமாக)
0	38	(38)(0)	0.03	33
1	144	(144)(1)	0.16	162
2	342	(342)(2)	0.32	316
3	287	(287)(3)	0.31	309
4	164	(164)(4)	0.15	151
5	25	25(5)	0.03	29
மொத்தம்	1000	2470		1000

அட்டவணை 41இல் கொடுக்கப்பட்டுள்ள தலைகளின் எண்ணிக்கையையும் அலைவெண்களையும் பயன்படுத்தி கூட்டுச் சராசரி கண்டுபிடிக்கப்பட்டுள்ளது. அது 2.47 ஆகும் (2470/1000). அங்கு ஐந்து நாணயங்கள் சுண்டப்பட்டன; எனவே $n = 5$. சராசரி $= np = 2.47$. எனவே, $5p = 2.47$, $p = 0.494$. எனவே, $p(x) = {}^{5}c_{x}p^{x}q^{n-x} = {}^{5}c_{x}(0.494)^{x}(0.506)^{n-x}$. இதைப் பயன்படுத்தி $X = 0, 1,2,3,4,5$ ஆக வைத்து $p(x)$ கண்டுபிடிக்கப்பட்டு அவை அட்டவணை 41இல் கொடுக்கப் பட்டுள்ளன. மேலும் கிடைத்துள்ள விபரங்களிலிருந்து எதிர்பார்க்கப்படும் அலைவெண்களும் கணிக்கப்பட்டுள்ளன. அவை சோதனை மூலம் கிடைத்துள்ள அலைவெண்களுக்குத் தோராயமாகச் சமமாக உள்ளன. சோதனை மூலம் கிடைத்த அலைவெண்களும் நிகழ்தகவைப் பயன்படுத்தி எதிர்பார்க்கப்படும் அலைவெண்களும் எந்த அளவுக்கு ஒத்துப் போகின்றன; அவை இரண்டுக்கும் உள்ள வித்தியாசங்கள் புள்ளியியல் முக்கியத்துவம் பெற்றதா, இல்லையா போன்ற விளக்கங்களை எடுகோள்களைச் சோதனை (Testing of Hypotheses) செய்யும் முறையை வைத்துத் தருவிக்கலாம்.

அட்டவணை 39இல் சமமான வாய்ப்பு கொண்ட தலையும் பூவும் உள்ள மூன்று நாணயங்களைச் சுண்டும்போது 0 முதல் 3 தலைகள் வரை வருவதற்கான நிகழ்தகவுகளும், அந்தப் பரவலின் கூட்டுச் சராசரி காணும் விதமும் கொடுக்கப்பட்டுள்ளன. மூன்று நாணயங்கள் சுண்டும்போது கிடைக்கின்ற நிகழ்தகவுகளுக்குள்ளே கொஞ்சம் வித்தியாசங்கள் இருக்கின்றன, வரைபடம் 32(அ)வில் உள்ளது போல். நாணயங்களின் எண்ணிக்கையைக் கூட்டக் கூட்ட இந்த வித்தியாசம் குறைந்து கொண்டே வரும். அதுபோல, நாணயங்களின் எண்ணிக்கையைக் கூட்டிக் கொண்டே செல்லும்போது, கிடைக்கின்ற நிகழ்வுகளின் எண்ணிக்கையும்

கூடிக்கொண்டே செல்லும். உதாரணத்திற்கு 0 முதல் 3 தலைகள் என்று வந்ததற்குப் பதிலாக 0 முதல் 1000 தலைகள் வரலாம் (1000 நாணயங்கள் சுண்டப்பட்டால்). அவ்வாறாக, வாய்ப்புக்களின் எண்ணிக்கை கூடும்போதும், நிகழ்தகவுகளுக்கு இடையே உள்ள வித்தியாசங்கள் குறையும் போதும் புள்ளிகள் நெருக்கமாக வந்து வரைபடம் 32(ஆ)வில் உள்ளதுபோல் காட்சியளிக்கும். இவ்வாறாக, ஒரு தனித்த நிகழ்வுப் பரவல் (discrete probability distribution) ஒரு தொடர் நிகழ்வுப் பரவலாக (Continuous probability distribution) மாறிவரும்; அடர்த்தி கூடிவரும்.

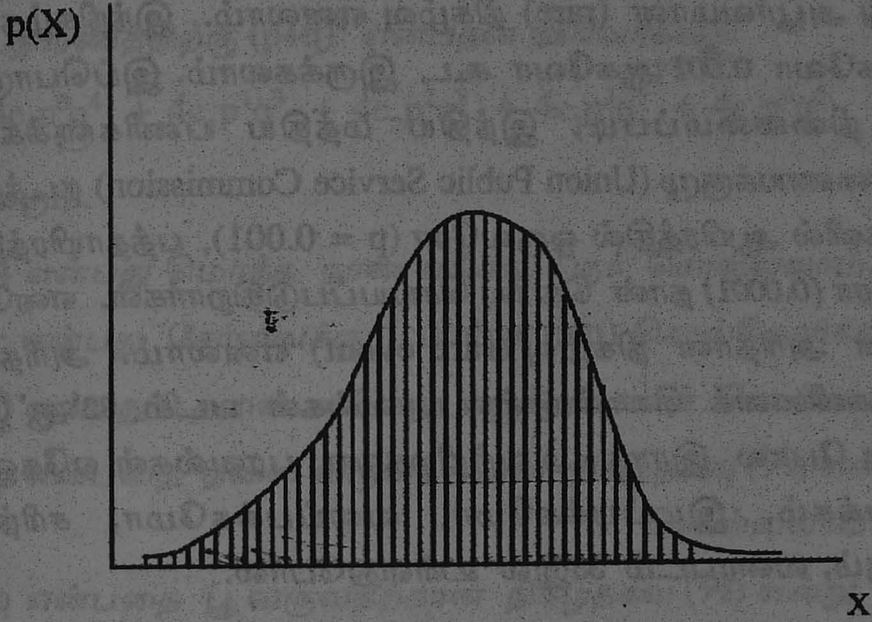
வரைபடம் - 32 (அ)



தலைகளின் எண்ணிக்கை

(அட்டவணை 39இல் உள்ள விபரங்கள்)

வரைபடம் - 32 (ஆ)

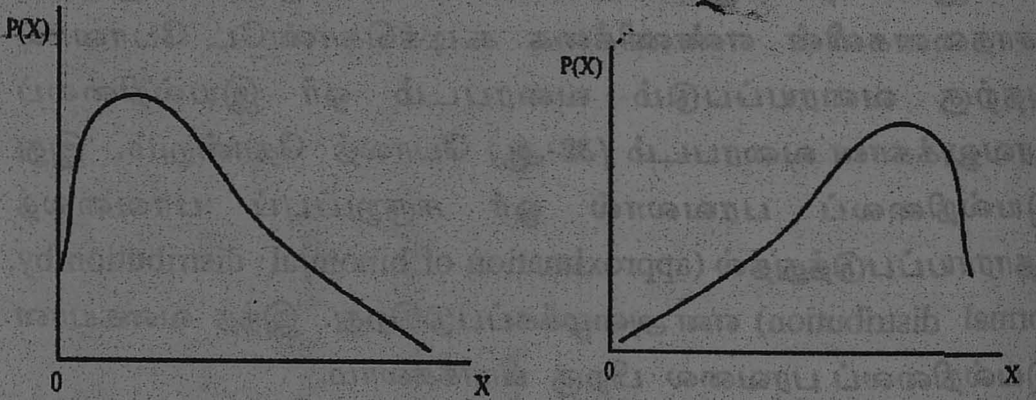


இவ்வாறாக இரண்டு சமவாய்ப்பு நிகழ்வுகள் இருந்து சோதனைகளின் எண்ணிக்கை கூடிக்கொண்டே போனால் அதற்கு வரையப்படும் வரைபடம் ஓர் இயல்நிலைப் பரவலுக்கான வரைபடம் (32-ஆ) போலத் தோன்றும். இது இயல்நிலைப் பரவலால் ஓர் ஈருறுப்புப் பரவலைத் தோராயப்படுத்துதல் (approximation of binomial distribution by normal distribution) என அழைக்கப்படுகிறது. இந்த வகையான இயல்நிலைப் பரவலை பிறகு பார்க்கலாம்.

மேலே கூறியதுபோல் சோதனைகளின் எண்ணிக்கை அதிகமாக இருந்து, இரண்டு நிகழ்வுகளும் சமவாய்ப்பு நிகழ்வுகளாக இல்லாதிருந்தால், ஈருறுப்புப் பரவல் பாய்ஸான் பரவலாக மாறிவிடும். உதாரணத்திற்கு, போட்டித் தேர்வுகளில் வெற்றி பெறுவது மிகவும் கடினம் என மாணவர்கள் நினைக்கிறார்கள் என்றால், அவற்றில் ஒரு மாணவர் தேர்ச்சி பெறுவதற்கான நிகழ்தகவு மிகவும் குறைவாக 0க்கு (பூஜ்யம்)

அருகில் (எடுத்துக்காட்டாக 0.1 என்று) இருக்கும் என்று சொல்லலாம். அப்படியானால், அந்த நிகழ்வை அரிதான அல்லது அபூர்வமான (rare) நிகழ்வு எனலாம். இந்நிகழ்தகவு 0.01 ஆகவோ 0.001 ஆகவோ கூட இருக்கலாம். இப்பொழுது உள்ள நிலைமைப்படி, இந்திய மத்திய பணிகளுக்கான தேர்வாணையக்குழு (Union Public Service Commission) நடத்தும் தேர்வுகளில் ஆயிரத்தில் ஒருவரோ ($p = 0.001$), பத்தாயிரத்தில் ஒருவரோ (0.0001) தான் தெரிவு செய்யப்படுகிறார்கள். எனவே, இதனை அரிதான நிகழ்வு (rare event) எனலாம். அரிதான நிகழ்வுகளினைக் கொண்டுள்ள பரவல்கள் படம் 32'ஆ' இல் உள்ளது போல இராது; அவற்றிற்கான பரவல்கள் ஏதேனும் ஒரு பக்கம், இடப்பக்கமோ, வலப்பக்கமோ, சரிந்தே இருக்கும், வரைபடம் 33இல் உள்ளதுபோல.

வரைபடம் - 33



எல்லாவிதமான பரவல்களின் போக்குகளையும் (Pattern or behaviour) அவற்றின் சராசரி பாதிக்கின்றது என்று சொல்லலாம்.

ஈருறுப்பின் விரிவாக்கமும் (Binomial expansion)
ஈருறுப்புக் கெழுக்களும் (Binomial coefficients)

ஒரு பரிசோதனை பல தடவைகள் (n) நடத்தப்பட்டு ஒவ்வொரு நிகழ்வுக்கும் எந்தளவு நிகழ்தகவு இருக்கின்றது

என்பதைக் கண்டறியும் முறையிலிருந்து ஈருறுப்பு விரிவாக்கமும் ஈருறுப்புக் கெழுக்களும் பெறப்படுகின்றன.

உதாரணத்திற்கு $(p+q)^4$ என்பதன் விரிவாக்கம்:

$$4c_0p^0q^4 + 4c_1p^1q^3 + 4c_2p^2q^2 + 4c_3p^3q^1 + 4c_4p^4q^0$$

என்பதாகும்.

இதில் 4 என்பது மொத்த நாணயங்களுக்குச் சொல்லலாம்.

c என்பது சேர்வைக்கு (combination) சொல்கிறார்கள்.

0 என்பது 0 தலை அல்லது 4 பூக்கள் எனச் சொல்லலாம்

p என்பதை தலை வருவதற்கான நிகழ்தகவு (½) என்று சொல்லலாம்.

q என்பதை பூ வருவதற்கான நிகழ்தகவு (½) என்று சொல்லலாம்.

இங்கு நான்கு நாணயங்கள் உள்ளதால், ஐந்து வெவ்வேறு விளைவுகளுக்கு வாய்ப்பு உண்டு. அதாவது *0* தலை *4* பூக்கள் *1* தலை *3* பூக்கள், *2* தலைகள் *2* பூக்கள், *3* தலைகள் *1* பூ, அல்லது *4* தலைகள் *0* பூ.

மூன்று நாணயங்களாக இருந்திருந்தால் மொத்த விளைவுகள் *4* ஆக இருக்கும். எனவே மொத்த விளைவுகளின் எண்ணிக்கை n+1 என்று சொல்லலாம். எனவே, ஈருறுப்பு விரிவாக்கத்தில் உள்ள பகுதிகள் (terms) n+1 ஆக இருக்கும்.

இங்கே கொடுக்கப்பட்டுள்ள ஈருறுப்பு விரிவாக்கத்தில் கெழுக்கள் என்பது $4c_0$, $4c_1$, $4c_2$, $4c_3$, $4c_4$ ஆகியவை ஆகும். $4c_0 = 1$, $4c_1 = 4$, $4c_2 = 6$, $4c_3 = 4$, $4c_4 = 1$ என்று பாஸ்கலுடைய முக்கோணத்திலிருந்தும் (PASCAL'S TRIANGLE) தெரிந்து கொள்ளலாம்.

பாஸ்கலுடைய முக்கோணம்

அட்டவணை 42 : பாஸ்கலுடைய முக்கோணம்

										1										
									1		1									
								1		2		1								
							1		3		3		1							
						1		4		6		4		1						
					1		5		10		10		5		1					
				1		6		15		20		15		6		1				
			1		7		21		35		35		21		7		1			
		1		8		28		56		70		56		28		8		1		
	1		9		36		84		126		126		84		36		9		1	
1		10		45		120		210		252		210		120		45		10		1

பாஸ்கலுடைய முக்கோணத்தை மனப்பாடம் செய்யத் தேவையில்லை. நேர் மேலேயுள்ள இரண்டு எண்களைக் கூட்டினால் கீழே உள்ள எண் வரும். நான்கு நாணயங்களாக இருந்தால் ஐந்தாவது வரிசையைப் பார்த்தால் தேவையான ஈருறுப்புக் கெழுக்கள் கிடைக்கும். ஐந்து நாணயங்களாக இருந்தால் கெழுக்களுக்கு ஆறாவது வரிசையைப் பார்க்கலாம். இந்தக் கெழுக்களை உற்றுக் கவனித்தால் அவை எவ்விதம் கூடுகின்றனவோ அதே மாதிரி குறைந்து கொண்டும் செல்வது தெரியும். நடுவில் வருகின்ற கெழு மிகப் பெரியதாக இருக்கும்.

இதனை முகடு (MODE) எனலாம். சோதனைக்கு உட்படுத்தப்படும் நாணயங்களின் எண்ணிக்கை இரட்டைப் படையாக (உ.ம். 4) இருந்தால், அங்கு விளைவுகள் அல்லது பகுதிகள் (binomial terms) ஒற்றைப்படை எண்ணில் (உ.ம்.5) இருக்கும். அப்படிப்பட்ட பரவலில் ஒரே ஒரு முகடுதான் இருக்கும். அதனை ஒரு-முகட்டுப் பரவல் (Unimodal distribution) என்பார்கள். (கொடுக்கப்பட்டுள்ள உதாரணத்தில் முகடு = 6). மாறாக, சோதனைக்குட்படுத்தப்படும் நாணயங்களின் எண்ணிக்கை ஒற்றைப்படையாக இருந்தால், (உ.ம்.5) அங்கு விளைவுகளின் எண்ணிக்கை (n+1) இரட்டைப் படையாக இருக்கும். (உ.ம். 6). அப்படிப்பட்ட பரவலில் இரண்டு முகடுகள் இருக்கும். அதனை இருமுகட்டுப் பரவல் (bimodal distribution) என்பார்கள்.

ஒவ்வொரு பகுதியிலும் உள்ள அடுக்குகள் இரண்டையும் கூட்டினால் அது மொத்த எண்ணிக்கைக்குச் (n) சமமாக இருக்கும். கொடுக்கப்பட்டுள்ள உதாரணத்தில் *(0+4), (1+3), (2+2), (3+1), (4+0)* ஆக உள்ளன.

பாய்ஸான் பரவல் (Poisson Distribution)

ஈருறுப்புப் பரவலும் பாய்ஸான் பரவலும் தனித்த நிகழ்தகவுப் பரவல்கள் என்றும் இரண்டிலும் மிக அதிகமாக இரண்டு நிகழ்வுகளே இருக்கமுடியும் என்றும் முன்னரே விளக்கப்பட்டது. இவற்றில் இவை இரண்டும் வேறுபாடற்று அமைந்தாலும், இவற்றிற்கிடையே ஒரு வித்தியாசம் உள்ளது. மேலே சொன்ன குணங்களுடன், இரு நிகழ்வுகளும் சமவாய்ப்பு (அல்லது ஏறத்தாழ சமவாய்ப்பு) நிகழ்வுகளாக இருந்தால் அது ஈருறுப்புப் பரவல். மாறாக, அவ்விரு நிகழ்வுகளும் (p மற்றும் q) மிகவும் விலகியிருந்தால் அதனைப் பாய்ஸான் பரவலுக்குள் கொண்டு வரலாம். உதாரணத்திற்கு, 100 ரூபாய் பணத்தாள்கள் 100 எடுக்கும்போது அதில் ஒரு கள்ளப் பணத்தாள் இருக்கலாம். (அறிவியல்

வளர்ச்சி அனைவரையும் சென்று சேரும்போது இந்த எண்ணிக்கை கூடலாம்). அப்படி ஒரு கள்ளப்பணத்தாள் இருந்தால் p=0.01 என்றும் கூறலாம். எனவே, ஒரு கள்ளப் பணத்தாளைக் கண்டுபிடிக்கும் நிகழ்தகவு (p) = 0.01 என்றும் சரியான பணத்தாளைப் பெற நிகழ்தகவு (q) என்றும் கூறலாம். இவ்வாறாக $p \approx 0; q \approx 1$ இருக்கும் சூழ்நிலைகளில் pயை அரிதான (rare) நிகழ்வு என்கிறார்கள். மாறாக $q \approx 0; p \approx 1$ஆக இருக்கும்போது qயை அரிதான நிகழ்வு என்கிறார்கள். முர்ரே ஆர். ஸ்பீஜெல் (MURRAY R. SPIEGEL) தன் புத்தகத்தில் (Schaum's Outline of Theory and Problems of Statistics, McGraw - Hill International Book Company, Singapure, 1961, p. 124) 50 முறைகள் சோதனை நடத்தப்பட்ட ஒரு பரவலின் சராசரி 5க்கும் குறைவாக இருந்தால், அதற்கான நிகழ்வினை அரிதான நிகழ்வு என்கிறார். அப்படியிருக்கும் நிலையில் அது பாய்ஸான் பரவலெனக் கருதலாம்.

இரு நிகழ்வுகளில் ஏதேனும் ஒன்று நிகழ்வதற்கான நிகழ்தகவு 0க்கு மிக அருகிலோ ஒன்றுக்கு (1) மிக அருகிலோ இருந்தால் சோதனைகளின் எண்ணிக்கை (n) அதிகமாக இருந்தாலும் ஈருறுப்புப் பரவலின் தோராயமாக இயல்நிலைப் பரவலைப் பயன்படுத்தமுடியாது. சோதனைகளின் எண்ணிக்கை (n) அதிகமாக இருந்து, நிகழ்வினில் ஒன்றின் நிகழ்தகவு (P) மிகக் குறைவாக இருந்தாலும் ஈருறுப்பு நிகழ்தகவு காண்பது கடினமான பணியாகும். இச்சூழலில் பாய்ஸான் பரவலைப் பயன்படுத்தலாமென ஃபிரெஞ்சு (FRENCH) கணிதமேதையான பாய்ஸான் (S.D.POISSON : 1781-1840) சுட்டிக்காட்டியுள்ளார்.

ஒரு பெட்டியில் 99 பச்சைப் பந்துகளும் 1 வெள்ளைப் பந்தும் இருந்து, சமவாய்ப்பு முறைப்படி 1 பந்தை எடுக்கும்போது பச்சைப்பந்து வருவதற்கான நிகழ்தகவு 0.99 என்றும் (q) வெள்ளைப் பந்து வருவதற்கான நிகழ்தகவு (p) 0.01

என்றும் தெரிந்ததே. இப்படிப்பட்ட சூழலில், பத்து பந்துகளைச் சமவாய்ப்பு முறைப்படி ஒரே சமயம் எடுத்தால் அவற்றில் சரியாக ஒன்று வெள்ளைப்பந்தாக இருப்பதற்கான நிகழ்தகவு என்ன?

இதில் p = 0.01; n = 10; np = 0.1; x = 1. இந்த விபரங்களைப் பயன்படுத்தி

$$p(X) = \frac{e^{-\lambda}\lambda^{x}}{X!}$$

என்ற சூத்திரத்தைப் பயன்படுத்தி தேவையான பதிலைப் பெறலாம். இதில் $\lambda = 0.1$. e என்பது மாறிலி. அதன் மதிப்பு (constant) 2.71828 ஆகும்.

$$p(X) = \frac{2.71828^{-0.1}0.1^{1}}{1!}$$ (இது ஒரு வெள்ளைப்பந்து இருப்பதற்கான நிகழ்தகவு ஆகும்)

மற்றொரு கணக்கையும் எடுத்துச் செய்து பார்க்கலாம். ஒரு பல்கலைக்கழகத்தில் பயிலும் மாணவர்களில் 10 விழுக்காடு மாணவர்களுக்கே அவர்களின் படிப்புக்கு ஏற்ற வேலை கிடைக்கின்றதென்று வைத்துக் கொள்வோம். அந்தப் பல்கலைக்கழகத்திலிருந்து 10 மாணவர்களைச் சமவாய்ப்பு முறையில் தேர்ந்தெடுத்தோமேயானால், அவர்களில் சரியாக இரண்டு பேருக்கு சரியான வேலை கிடைக்க நிகழ்தகவு என்னவாக இருக்கும்?

இதில் $\lambda = np = 10\ (0.1) = 1$; $e = 2.71828$; $x = 2$

$$p(2\ \text{மாணவர்கள்}) = \frac{e^{-\lambda}\lambda^{x}}{X!} = \frac{2.71828^{-1}(1)^{2}}{2 \times 1} = 0.1839\ (\text{or})\ 0.184$$

இதே கணக்கை ஈருறுப்புப் பரவல் முறையில் செய்தோமேயானால்,

$$p(2\ \text{மாணவர்கள்}) = 10c_{2}p^{2}q^{8}$$

$$= \frac{10 \times 9}{1 \times 2} (0.1)^2 (0.9)^8 = 0.19$$

அனேகமாக இரு முறைகளிலும் நிகழ்தகவுகள் ஏறத்தாழச் சமமாகவே உள்ளன.

இதில் 50 மாணவர்கள் சமவாய்ப்பு முறையில் எடுக்கப்பட்டிருந்தால், பாய்ஸான் பரவலில்,

$\lambda = np = 50\ (0.1) = 5$

$x = 2$

p(2 மாணவர்கள்) $= \frac{2.71828^{-5} 5^2}{2}$

ஈருறுப்புப் பரவலில்,

$50c_2p^2q^{48}$

$= \frac{50 \times 49}{1 \times 2} (0.1)^2 (0.9)^{48}$ இது மிகவும் சிரமமாக இருக்கும்.

அதே கணக்கில் 50 மாணவர்கள் சமவாய்ப்பு முறையில் எடுக்கப்பட்டிருந்து, 10 மாணவர்களுக்குப் பொருத்தமான வேலை கிடைப்பதற்கு நிகழ்தகவு என்ன என்று கேட்டிருந்தால், பாய்ஸான் முறைப்படி,

$\lambda = np = 50\ (0.1) = 5; \quad x = 10$

p(10 மாணவர்கள்) $= \frac{2.71828^{-5} 5^{10}}{10!}$

பாய்ஸான் முறைப்படி பதில் கிடைப்பதை விட ஈருறுப்புப் பரவலின்படி இன்னும் சிரமமாக இருக்கும்.

ஈருறுப்புப் பரவல் மூலம் p(10 மாணவர்கள்) $= 50c_{10}p^{10}q^{40}$

$$= \frac{50 \times 49 \times \ldots \times 41}{1 \times 2 \times \ldots \times 10} (0.1)^{10} (0.9)^{40}$$

இவ்வாறாக nம் xம் அதிகமாக இருக்கும்போது, ஈருறுப்புப் பரவலையும் பாய்ஸான் பரவலையும் விட இயல்நிலைப் பரவல் மூலம் விடை காண்பது எளிதாக இருக்கும்.

கீழ்வரும் அட்டவணை 43இல் ஒரு நகரில் நடந்த விபத்துக்களின் விபரங்கள் தரப்பட்டுள்ளன. அவற்றைப் பயன்படுத்தி பாய்ஸான் பரவலைத் தருவிக்கலாம்.

அட்டவணை - 43

விபத்துக்களின் எண்ணிக்கை	விபத்து நடந்த நாட்கள்	fx	நிகழ்தகவு	எதிர்பார்க்கப் படும் நாட்கள்
0	21	0	0.41	20.3
1	18	18	0.37	18.3
2	7	14	0.16	8.2
3	3	9	0.05	2.5
4	1	4	0.01	0.6
Total	**50**	**45**	**1.00**	

$$\text{இதன் சராசரி} = \frac{45}{50} = 0.90$$

$p(X) = \frac{e^{-\lambda}\lambda^{x}}{X!}$ இந்தச் சூத்திரத்தைப் பயன்படுத்தி *0, 1, 2, 3, 4* ஆகிய விபத்துக்கள் நடப்பதற்கான நிகழ்தகவுகளைக் கணக்கிடலாம். அவையும் அட்டவணை 43இல் கொடுக்கப்பட்டுள்ளன. உண்மையிலேயே விபத்து நடந்த நாட்களின் எண்ணிக்கைக்கும் விபத்து நடக்க எதிர்பார்க்கப்படும் நாட்களின் எண்ணிக்கைக்கும் வித்தியாசம் மிகக் குறைவாகவே உள்ளது. இந்த வித்தியாசம் புள்ளியியல் முக்கியத்துவம் பெற்றதா என்று சோதித்துப் பார்த்தால் முக்கியத்துவம் பெற்றில்லை என்ற முடிவு வருவதற்கே வாய்ப்புக்கள் அதிகம்.

அட்டவணை - 44
பாய்ஸான் பரவலின் பண்புகள்

கூட்டுச் சராசரி	(Mean)	$\mu = \lambda$ (lamda)
மாறுபாடு	(Variance)	$\sigma^2 = \lambda$
திட்டவிலக்கம்	(Standard deviation)	$\sigma = \sqrt{\lambda}$
கோட்டத்தின் அசைவுக்கெழு	(Moment coefficient of Skewness)	$\alpha_3 = 1 \div \sqrt{\lambda}$
தட்டைத் தன்மையின் அசைவுக்கெழு	(Moment coefficient of Kurtosis)	$\alpha_4 = 3 + (1 \div \lambda)$

இயல்நிலைப் பரவல் (Normal Distribution)

இயல்நிலைப் பரவலை முறைப்படுத்தியவர்களுள் முக்கியமானவர்கள் ஆப்ரஹாம் டி மோய்வ்ரே (ABRAHAM De MOIVRE : 1667 - 1754), பியரே எஸ்.லேப்லேஸ் (PIERRE S. LAPLACE : 1749-1827), கார்ல் காஸ் (KARL GAUSS : 1777 - 1855) ஆவர். டி மோய்வ்ரே முதன்முதலில் இயல்நிலைப் பரவலைப் பற்றி விவரித்திருந்த போதும், அவருடைய அந்த அரும்பணி கவனிப்பாரற்றுப் போயிற்று. அவருக்குப் பின்னர் வந்த காஸ் என்பவரின் பங்களிப்பு கணிதமேதைகளின் மத்தியில் ஒரு சிறப்பான இடத்தைப் பிடித்தது. எனவேதான் இயல்நிலைப் பரவல் காஸியன் பரவல் என்றும் அழைக்கப் பெறுகிறது. இயல்நிலைப் பரவலின் வரைபடம் இயல்நிலை வளைகோடு (Normal curve) என்றும், காஸியன் வளைகோடு (Gaussian curve) என்றும், இயல்நிலை நிகழ்தகவு வளைகோடு (Normal Probability Curve) என்றும் பிழைகாண உதவும் இயல்நிலை வளைகோடு (Normal Curve of Error) என்றும் அழைக்கப்படுகின்றது.

இயல்நிலை வளைகோட்டை வரைய உதவும் சமன்பாடு

$$Y = \left[\frac{N}{\sigma\sqrt{2\pi}}\right] [e^{w}]$$

$$W = \frac{-(x-\mu)^2}{2\sigma^2}$$

இதில் μ என்பது (Mu) பரவலின் கூட்டுச்சராசரி

σ என்பது (Sigma) பரவலின் திட்டவிலக்கம்

N என்பது அலைவெண்களின் மொத்தம்

நிகழ்தகவு அடர்த்தி சார்பு $Y = \left[\frac{1}{\sigma\sqrt{2\pi}}\right] [e^{w}]$

$$W = \frac{-(x-\mu)^2}{2\sigma^2}$$

இதில் X என்ற தொடர் சமவாய்ப்பு மாறியினை நிலைப்படுத்த $\frac{x-\mu}{\sigma}$ என்ற சூத்திரம் பயன்படுத்தப்படுகிறது. இதில் உள்ள X என்பது எதுவேண்டுமானாலும் இருக்கலாம்; எந்த அலகினாலும் அளவிடப்படலாம். இவற்றினால் விளையும் வேறுபாடுகளை அகற்றுவதற்காகவே X நிலைப்படுத்தப் படுகிறது (standardised).

எடுத்துக்காட்டாக, மாணவர்களின் படிப்புத்திறமையை (கல்வித்திறமை என்பதும் படிப்புத்திறமை என்பதும் வித்தியாசமாகலாம்) அளவிட ஒன்றிற்குமேற்பட்ட நிறுவனங்கள் தேர்வுகள் நடத்தலாம்; ஒன்றிற்கு மேற்பட்ட பாடவகைகள் இருக்கலாம்; மொத்த மதிப்பெண்கள் வித்தியாசமாக இருக்கலாம் (அந்த நிறுவனங்கள் மொத்த மதிப்பெண்களை 500 என்றோ, 600 என்றோ, 1000 என்றோ 1200 என்றோ வைத்துக் கொள்ளலாம்). சில நிறுவனங்கள், மாணவர்களை மகிழ்விப்பதற்காகவும், ஊக்குவிப்பதற்காகவும்

அதிக மதிப்பெண்களை அளிக்கும் விதம் வினாக்களை அமைப்பதையும் விடைத்தாள்களைத் திருத்துவதையும் வைத்துக் கொள்ளலாம். சில நிறுவனங்கள், மாணவர்களின் மனவலிமையைக் கூட்டுவதற்காக, இலகுவாக அதிக மதிப்பெண்கள் பெறமுடியாத வகையில் தேர்வு முறைகளை அமைக்கலாம். இச்சூழலில், வெவ்வேறு நிறுவனங்கள் அளித்துள்ள மதிப்பெண்களை அப்படியே எடுத்துக்கொண்டு மாணவர்களின் படிப்புத்திறமையை மதிப்பிடுவது சரியாக இருக்க முடியாது. எனவே, மாணவர்கள் பெற்றுள்ள மதிப்பெண்களை நிலைப்படுத்திவிட்டு (standardisation) ஒப்பிடுவது சரியாக இருக்கும்.

எளிமையான தேர்வு முறைகளைக் கடைப்பிடிக்கும் நிறுவனத்தில் (அ) மாணவர்கள் பெற்ற மதிப்பெண்களின் சராசரி 60 ஆகவும் திட்டவிலக்கம் 20 ஆகவும் இருக்கின்றதென்று கொள்வோம். கடினமான தேர்வு முறைகளைக் கடைப்பிடிக்கும் 'ஆ' என்ற நிறுவனத்தின் மாணவர்கள் பெற்ற மதிப்பெண்களின் சராசரி 50 ஆகவும் திட்டவிலக்கம் 10 ஆகவும் இருக்கிறதென்று கொள்வோம். இச்சூழலில், 'அ' நிறுவனம் நடத்திய தேர்வில் ஒரு மாணவர் (எ) 80 விழுக்காடு மதிப்பெண்களும், 'ஆ' நிறுவனம் நடத்திய தேர்வில் ஒரு மாணவர் (ஏ) 70 மதிப்பெண்களும் பெற்றுள்ளார் என்று வைத்துக்கொண்டு 80 மதிப்பெண்கள் பெற்ற 'எ' என்பவர் 'ஏ' என்பவரை விடப் படிப்பறிவில் சிறந்தவர் என முடிவு செய்வது சரியா என்று இங்கு பார்க்கலாம்.

மாணவர் 'எ' பெற்ற மதிப்பெண்களை நிலைப்படுத்தினால்,

$$z = \frac{80 - 60}{20} = \frac{20}{20} = 1$$ என்று வருகிறது.

மாணவர் 'ஏ' பெற்ற மதிப்பெண்களை நிலைப்படுத்தினால்,

$$z \frac{70 - 50}{10} = \frac{20}{10} = 2$$ என்று வருகிறது.

நிலைப்படுத்தப்பட்ட மதிப்பெண்களிலிருந்து மாணவர் 'ஏ' மாணவர் 'எ'யை விடப் படிப்பறிவில் சிறந்தவர் என்று முடிவு செய்வது சரியாகத் தெரிகிறது. இவ்வாறாக நிலைப்படுத்துதல் பயனளிக்கிறது.

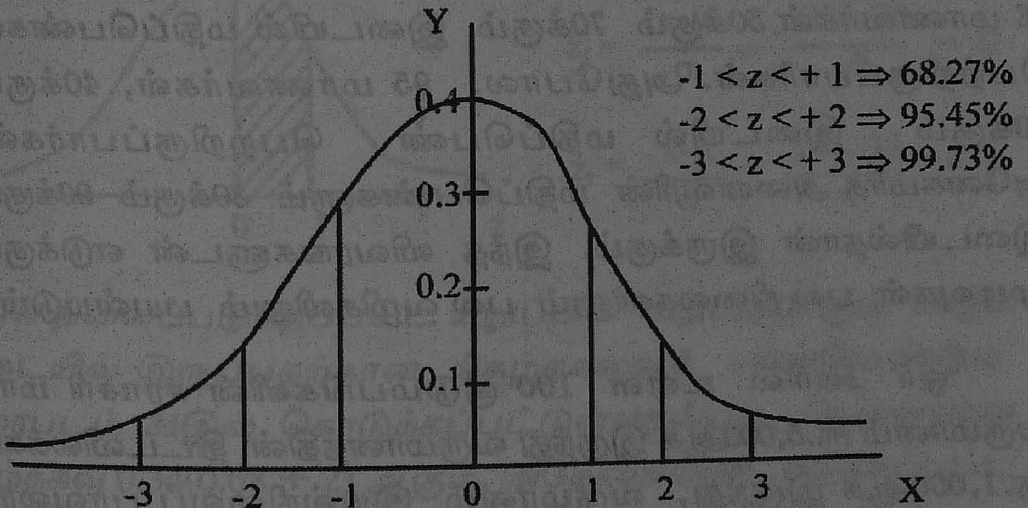
X நிலைப்படுத்தப்பட்ட பின்னர், நிகழ்தகவு அடர்த்தி சார்பு

$$Y = \left[\frac{1}{\sqrt{2\pi}}\right] [e^{w}]$$

$$W = \frac{1}{2} z^2$$

நிலைப்படுத்தப்பட்ட Xயை z எனக் குறிக்கிறார்கள். இந்த zஇன் கூட்டுச்சராசரி 0 ஆகவும், மாறுபாடு (variance) 1 ஆகவும் இருக்கும்.

வரைபடம் 34



நிலைப்படுத்தப்பட்ட X, அதாவது z ன் மதிப்பு -1 முதல் +1 வரை இருக்கும்போது அதற்கிடையே இருக்கும் பரப்பு 68.27 சதவீதமாகும். அதாவது, இயல்நிலை வளைகோடு மூடியிருக்கும் பரப்பளவு 100 என்று கொண்டால், 68.27 சதவீதப் பரப்பளவு -1 முதல் +1 வரையுள்ள தூரத்திற்குள் இருக்கும். இதை நிகழ்தகவில் சொல்வதென்றால் 0.6827 ஆகும். வேறுவிதமாகக் கூறினால், 100 மாணவர்களின் மதிப்பெண்கள் இயல்நிலைப் பரவலைக் கொண்டிருந்தால், நிலைப்படுத்தப்பட்ட -1க்கும் +1க்கும் இடையில் 68 மாணவர்கள் மதிப்பெண்கள் பெற்றிருப்பார்கள். ஒரு முழுமையிலிருந்து 100 சமவாய்ப்புக் கூறுகள் எடுத்தால், அவற்றில் 68 கூறுகள் (samples) நிலைப்படுத்தப்பட்ட -1க்கும் +1க்கும் இடையில் சராசரியைக் கொண்டு இருக்கும்.

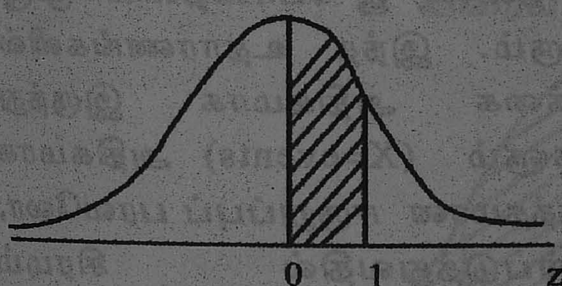
உதாரணத்திற்கு, ஒரு கல்லூரியில் உள்ள மாணவர்களின் சராசரி மதிப்பெண் 60 ஆகவும் திட்டவிலக்கம் 10 ஆகவும் இருந்து மதிப்பெண் இயல்நிலைப் பரவலாக இருந்தால், 50க்கும் 70க்கும் இடையில் மதிப்பெண் பெற்ற மாணவர்கள் 68 சதவீதம் ஆகும். அந்தக் கல்லூரியில் இருந்து 100 மாணவர்களைச் சமவாய்ப்பு கூறுகளாக எடுத்தோமேயானால் 68 மாணவர்கள் 50க்கும் 70க்கும் இடையில் மதிப்பெண்கள் பெற்றிருப்பார்கள். அதுபோல, 95 மாணவர்கள், 40க்கும் 80க்கும் இடையில் மதிப்பெண் பெற்றிருப்பார்கள். அனேகமாக அனைவரின் மதிப்பெண்களும் 30க்கும் 90க்கும் இடையில்தான் இருக்கும். இந்த விவரங்களுடன் எடுக்கும் முடிவுகள் பல நிலைகளிலும் பல வழிகளிலும் பயன்படும்.

ஓர் ஊரில் உள்ள 100 குடும்பங்களின் சராசரி மாத வருமானம் ரூ.5,000ஆக இருந்து வருமானத்தின் திட்டவிலக்கம் ரூ.1,000ஆக இருந்து, வருமானம் இயல்நிலைப் பரவலாக இருந்தால், அங்கு 68 வீடுகளின் வருமானம் ரூ.4,000க்கும்

ரூ.6,000க்கும் இடையில் இருக்கும். பதினாறு வீடுகளின் மாத வருமானம் ரூ.4,000க்குக் கீழ் இருக்கும். பதினாறு வீடுகளின் மாத வருமானம் ரூ.6,000க்கு மேல் இருக்கும். இதுபோல் இன்னும் பல விபரங்களை இயல்நிலைப் பரவலைப் பற்றிய புரிதல் தர வாய்ப்பிருக்கிறது. இந்த விபரங்கள் பொருளியல் கொள்கை வகுப்பவர்களுக்கு மிகவும் பயனுள்ளதாக இருக்கும். ஒரு மாநிலத்தில், மாவட்டத்தில், வட்டத்தில் அல்லது கிராமத்தில் எத்தனை வீடுகள் ஏழ்மைக்குள் தள்ளப்பட்டிருப்பார்கள் என்பதை டெல்லியிருந்தே தெரிந்து கொள்ளலாம்; சராசரி, திட்டவிலக்கம், வீடுகளின் எண்ணிக்கை ஆகியவை மட்டும் தெரிந்தால் போதும்.

ஓர் இயல்நிலைப் பரவலின் கூட்டுச் சராசரி 20 ஆகவும் திட்டவிலக்கம் 4 ஆகவும் இருக்கின்றன என்று கொள்வோம். அப்படியானால் அந்தப் பரவலிலிருந்து ஓர் உறுப்பை (element) சமவாய்ப்பு முறையில் எடுத்தால் அது 20க்கும் 24க்கும் இடையில் இருப்பதற்கான நிகழ்தகவு என்ன?

வரைபடம் - 35

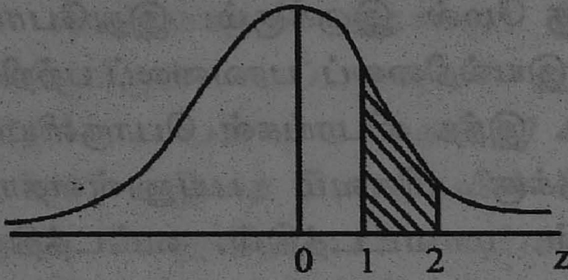


$$z_1 = \frac{x-\mu}{\sigma} = \frac{20-20}{4} = 0$$

$$z_2 = \frac{x-\mu}{\sigma} = \frac{24-20}{4} = 1$$

இதில் நிலைப்படுத்தப்பட்ட உறுப்பின் மதிப்பு 0க்கும் 1க்கும் இடையில் இருப்பதற்கான நிகழ்தகவைக் காணவேண்டும். வரைபடம் 34இல் கொடுக்கப்பட்டுள்ளதன்படி, தேவையான நிகழ்தகவு 0.34(0.68 ÷ 2) ஆகும். அதேகணக்கில், எடுக்கப்பட்ட உறுப்பு 24க்கும் 28க்கும் இடையில் இருப்பதற்கான நிகழ்தகவு எவ்வளவு?

வரைபடம் - 36



$$z_1 = \frac{x-\mu}{\sigma} = \frac{24-20}{4} = 1$$

$$z_2 = \frac{x-\mu}{\sigma} = \frac{28-20}{4} = 2$$

z = 0 முதல் z = 1 வரையிலான நிகழ்தகவு (பரப்பளவு) *0.34* ஆகும். z = 0 முதல் z = 2 வரையிலான நிகழ்தகவு 0.475 ஆகும். எனவே z = 1க்கும் z = 2க்கும் இடையில் உள்ள பரப்பளவு 47.5 - 34.0 = 13.5 விழுக்காடு. எனவே, நிகழ்தகவு *0.135* ஆகும்.

இதுவரையில் எடுக்கப்பட்ட உதாரணமாகிய மதிப்பெண் என்பது தொடர் மாறியாகும் (Continuous variable). சில சமயம் தனித்த மாறிகள் (discrete variables) இருக்கலாம். உதாரணத்திற்கு, முன்னர் பயன்படுத்தப்பட்ட உதாரணங்களாகிய கள்ளப் பணத்தாள், வீடுகள், மாணவர்கள், விபத்துக்கள், நாணயத்தின் தலை, பூ போன்றவை முழு எண்களாகத் தான் வரும். இந்த உதாரணங்களில் சோதனைகளின் எண்ணிக்கை அதிகமாக இருந்து தேவைப்படுகின்ற நிகழ்வுகளும் (X=events) அதிகமாக இருந்தால், அதற்குப் பொருத்தமான ஈருறுப்புப் பரவலோ, பாய்ஸான் பரவலோ பயன்படுத்துவதில் சிரமம் அதிகமாக இருக்கும். இச்சூழலில், இயல்நிலைப் பரவலைப் பயன்படுத்தலாம். ஆனால், இயல்நிலைப் பரவல் தொடர்மாறிக்குத்தான் உருவாக்கப்பட்டது. எனவே தனித்த மாறிகளைத் தொடர்மாறிகளாக மாற்றத் 'தொடர்ச்சித் திருத்தம்' (continuity correction) செய்ய வேண்டியுள்ளது.

உதாரணத்திற்கு, 100 நாணயங்களைக் குலுக்கிப் போடும்போது 45க்கும் 55க்கும் இடையில் தலைகளின் எண்ணிக்கை வருவதற்கு நிகழ்தகவு எவ்வளவு என்னும் வினாவினை எடுத்துக் கொள்வோம். இது ஈருறுப்புப் பரவலுக்குப் பொருத்தமானது. இதனை,

$$[{}^{100}c_{45}p^{45}q^{55}] + [{}^{100}c_{46}p^{46}q^{54}] + \ldots + [{}^{100}c_{55}p^{55}q^{45}]$$

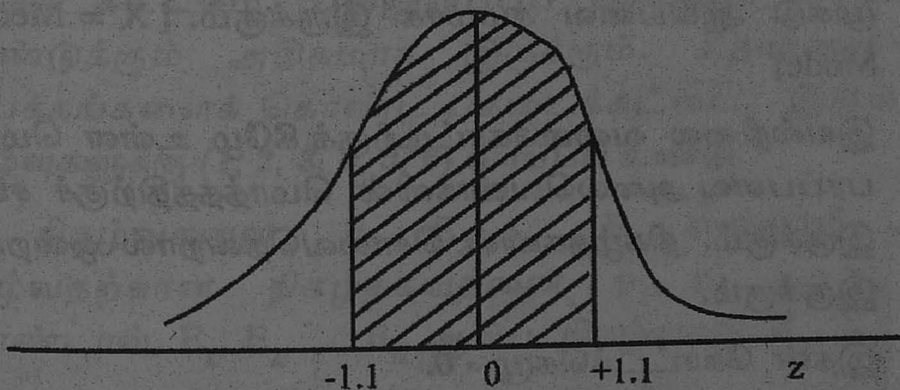
என்று கண்டுபிடிக்க வேண்டும். இது அதிக நேரமும் எடுக்கும்; சிரமமாகவும் இருக்கும். எனவே, இதை 'தொடர்ச்சித் திருத்தம்' (Continuity correction) செய்து இயல்நிலைப் பரவலாக்கி, எளிதாக (சரியான) விடை காணலாம்.

$$z_1 = \frac{44.5 - 50}{5} = \frac{-5.5}{5} = -1.1$$

$$z_2 = \frac{55.5 - 50}{5} = \frac{5.5}{5} = 1.1$$

வரைபடம் 37இல் நிழலாக்கப்பட்ட பரப்பளவை புள்ளியியல் அட்டவணையில் (Statistical Table) இருந்து மிக எளிதாகவும் சரியாகவும் கண்டுபிடித்து விடலாம்.

வரைபடம் - 37



0 முதல் +1.1 வரையிலான பரப்பளவு	= 0.3438
0 முதல் -1.1 வரையிலான பரப்பளவு	= 0.3438
மொத்தப் பரப்பளவு	= 0.6876

கேட்கப்பட்டுள்ள வினாவுக்குச் சரியான விடை 0.6876 ஆகும்.

இதிலிருந்து 100 நாணயங்களை 100 தடவைகள் குலுக்கினால் 68 அல்லது 69 தடவைகள் 45 முதல் 55 தலைகளுக்குள் வரும். இந்த முறையில் செய்வதனை 'ஈருறுப்புப் பரவலை இயல்நிலைப் பரவலுக்கு தோராயப்படுத்துதல்' எனலாம்.

இயல்நிலை வளைகோட்டின் தன்மைகள்

1. இயல்நிலை வளைகோடானது ஒரு முகட்டு (unimodal) மணி வடிவம் கொண்ட சமச்சீர் அணுகுகோடாகும் (asymptotic). இயல்நிலை வளைகோட்டின் இரு முனைகளுக்கும் X அச்சுக்கும் உள்ள தூரம் குறைந்து கொண்டே சென்று முனைகள் X அச்சைத் தொடாமலேயே இருக்கும்; ஆனால் தொடுவது போல் தோன்றும்.

2. இயல்நிலைப் பரவலின் கூட்டுச் சராசரி, இடைநிலை, முகடு ஆகியவை சமமாக இருக்கும். [$\overline{X}$ = Median = Mode]

3. இயல்நிலை வளைகோட்டிற்குக் கீழே உள்ள மொத்தப் பரப்பளவு அலைவெண்களின் மொத்தத்திற்குச் சமமாக இருக்கும். நிகழ்தகவில் சொல்வதென்றால் ஒன்றாக (1) இருக்கும்.

4. இதன் கோட்டக்கெழு = 0.

5. இதன் தட்டை அளவை = $\beta_2 = 3$.

6. மேல்கால்மானமும் கீழ்க்கால்மானமும் இடைநிலையிலிருந்து சமதூரத்தில் அமைந்திருக்கும்.

7. கால்மான விலக்கம் = $\frac{2}{3}\sigma$

8. சராசரி விலக்கம் = $\frac{4}{5}\sigma$

9. $\bar{X} \pm 1\sigma$ க்கிடையே ஏறத்தாழ 68.27 சதவீத இனங்களின் மதிப்புக்கள் இருக்கும். ஏறத்தாழ 95.45 சதவீத இனங்ளின் மதிப்புக்கள் $\bar{X} \pm 2\sigma$ க்கிடையே இருக்கும். ஏறத்தாழ 99.73 சதவீத இனங்களின் மதிப்புக்கள் $\bar{X} \pm 3\sigma$ க்கிடையே இருக்கும். இதனை வரைபடம் 34இல் காணலாம்.

மேலும் இரு பரவல்கள்

மேலே கூறப்பட்டுள்ள மூன்று பரவல்களுடன் இன்னும் இரண்டு வகைப் பரவல்கள் உள்ளன. அவற்றை இங்கு மிகச் சுருக்கமாகக் காணலாம். இவை இரண்டுமே தனித்த மாறிகளின் பரவல்கள்தான். இவற்றில் ஒன்று பல்லுறுப்புப் பரவல் (Multinomial distribution) மற்றொன்று அதிபெருக்குப் பரவல் (Hypergeometric distribution) ஆகும்.

பல்லுறுப்புப் பரவல்

பல்லுறுப்புப் பரவல் என்பது ஈருறுப்புப் பரவலின் அடுத்த கட்டமே. பல்லுறுப்புப் பரவலில் வாய்ப்புகள் இரண்டுக்கும் அதிகமாக இருக்கும். உதாரணம் ஆறு சமபக்கங்களைக் கொண்ட பகடைக்கட்டை. இதில் ஆறு நிகழ்வுகளுக்கு (1, 2, 3, 4, 5, 6) வாய்ப்பு உண்டு.

நிகழ்வுகளை $E_1, E_2, \ldots, E_K$ எனலாம். அவை நிகழ்வதற்கான நிகழ்தகவுதனை $p_1, p_2, \ldots, p_k$ எனக் கொண்டால், $E_1, E_2, \ldots, E_K$ ஆகிய நிகழ்வுகள், $X_1, X_2, \ldots, X_K$ தடவை நிகழ்வதற்கான நிகழ்வு : $\frac{N!}{X_1!\, X_2! \ldots X_k!}\ p_1^{X_1}, p_2^{X_2}, \ldots,$ $p_k^{X_K}$ ஆகும். இதில் $X_1 + X_2 + \ldots + X_K = N$.

பல்லுறுப்புப் பரவலின் விரிவாக்கம் : $(p_1 + p_2 + ... + p_K)^N$

உதாரணத்திற்கு, ஒரு பேதமற்ற பகடைக்காயை 12 தடவைகள் உருட்டினால், 1, 2, 3, 4, 5, 6 ஆகிய எண்கள் ஒவ்வொன்றும் சரியாக இரண்டு தடவைகள் மேற்புறத்தில் வருவதற்கான நிகழ்தகவு யாது? இதில் இரண்டிற்கும் மேற்பட்ட (ஆறு) நிகழ்வுகளுக்கு வாய்ப்பு இருப்பதால், இது ஒரு பல்லுறுப்புப் பரவலாகும்.

இந்த வினாவுக்கு விடை :

$$\frac{12!}{2!\,2!\,2!\,2!\,2!\,2!} \left(\frac{1}{6}\right)^2 \left(\frac{1}{6}\right)^2 \left(\frac{1}{6}\right)^2 \left(\frac{1}{6}\right)^2 \left(\frac{1}{6}\right)^2 \left(\frac{1}{6}\right)^2$$

$$= \frac{1925}{559872} = 0.00344$$

பல்லுறுப்புப் பரவலின் சராசரி, மாறுபாடு (variance) இன்னும் மற்ற விபரங்களுக்கு டேரோ யமனேயின் (TARO YAMANE) புள்ளியியல் : ஓர் அறிமுக ஆய்வு (Statistics : An Introductory Analysis) என்ற புத்தகத்தில் (Third Edition) 756ஆவது பக்கத்திலிருந்து படிக்கலாம்.

அதி பெருக்குப் பரவல் (Hypergeometric distribution)

ஒரு முழுமையிலிருந்து (populaiton) மாதிரிகள் (sample) எடுக்கும்போது முதலில் எடுத்த புள்ளிகளை திரும்ப வைக்காமல் (without replacement) அப்புறப்படுத்திவிட்டு மாதிரிகள் (samples) எடுத்தால், அப்போது அதிபெருக்குப் பரவல் பயன்படுத்தப்படுகிறது. எனவே, இவ்வகைப் பரவல் தரக்கட்டுப்பாடு புள்ளியியலில் (Quality Control Statistics) மிகவும் பயனுள்ளதாக இருக்கிறது. இப்பரவலை ஒரு பட்டியலாக லைபர்மேன் & ஓவன் (G.J.Lieberman and D.B.Owen, Tables of the Hypergeometric Distribution, Stanford University Press, 1960) ஆகியோர் தந்துள்ளார்கள்.

உதாரணத்திற்கு, ஒரு பெட்டியில் 60 சிவப்புப் பந்துகளும் 40 பச்சைப் பந்துகளும் உள்ளன. அவற்றில் இருந்து 10 பந்துகள் தேர்ந்தெடுக்கப்படுகின்றன. அவற்றில் 4 சிவப்புப் பந்துகளாக இருப்பதற்கு நிகழ்தகவு யாது? இந்தக் கணக்கை அதிபெருக்குப் பரவலைக் கொண்டு செய்யலாம்.

விடை $\frac{({}^{60}c_4)({}^{40}c_6)}{{}^{100}c_{10}}$ ஆகும்.

இன்னுமொரு எடுத்துக்காட்டை எடுத்துக் கொள்ளலாம். ஒரு வகுப்பில் ஆறு மாணவர்களும் நான்கு மாணவிகளும் உள்ளனர். அந்த வகுப்பிலிருந்து 5 பேர்கள் மாதிரியாகத் தேர்வு செய்யப்படுகின்றனர். அதில் 3 மாணவர்கள் இருப்பதற்கான நிகழ்தகவு யாது?

விடை $\frac{({}^{6}c_3)({}^{4}c_2)}{{}^{10}c_5}$ ஆகும்.

மீதமுள்ள மூன்று பரவல்களான χ^2 பரவல், t பரவல் F பரவல் ஆகியவைகளை எடுகோள்களைச் சோதனை செய்யும்போது காணலாம்.

புள்ளியியல் பட்டியல்கள்

ஈருறுப்பு, பாய்ஸான், இயல்நிலை, χ^2, t, F ஆகிய பரவல்களைக் கொண்ட கணக்குகளை எளிதாகச் செய்வதற்காகப் பல வகையான பட்டியல்கள் வந்துள்ளன. ஒவ்வொரு பரவலுக்கும் ஒன்றுக்கு மேற்பட்ட புள்ளியியல் பட்டியல்கள் சிறு வித்தியாசங்களுடன் வந்துள்ளன. அவற்றுள் சில இங்கு தரப்பட்டுள்ளன. அந்தப் பட்டியல்களைப் பயன்படுத்தும்போது ஒவ்வொரு பட்டியலுக்கும் பொருத்தமான பயன்படுத்தும் முறையை அறிந்து பயன்படுத்துதல் நலம் பயக்கும்.

௸

9. தீர்மானப்புள்ளியியல்
(INFERENTIAL STATISTICS)

பொதுவாக முழுமையைப் (Population or Universe) பற்றி அறிவதுதான் ஆய்வுகளின் நோக்கம். ஆனால், முழுமையுடைய மையப்போக்கு அளவுகளோ, சிதறல் அளவுகளோ, கோட்ட அளவுகளோ, தட்டைத்தன்மை அளவுகளோ எளிதாகக் கிடைப்பதில்லை; அவற்றைச் சேகரிப்பதும் சிரமம்; அதிக காலம், பணம் செலவிட வேண்டி வரலாம்; அவற்றைச் சேகரிக்கும்போது தவறுகளும் (mistakes) பிழைகளும் (errors) நிகழலாம். சோதித்து அறியப்பட வேண்டிய சூழ்நிலைகளில் சோதனைக்குப் பிறகு சிலவற்றின் பயன்பாடு குறைந்தோ பயன்பாடே இல்லாமலோ கூடப் போகலாம். உதாரணமாக, ஒரு நிறுவனம் உற்பத்தி செய்த எல்லா விளக்குகளும் எத்தனை மணி நேரம் ஒளிர்கின்றன என்று பார்ப்பதற்கான சோதனை செய்துவிட்டால், விற்பதற்கு விளக்குகள் இல்லாமல் போகலாம். எனவே இதுபோன்ற சூழ்நிலைகளில், ஒரு நிறுவனம் உற்பத்தி செய்த சில விளக்குகளை மட்டும் சோதித்துப் பார்த்துவிட்டு அந்த நிறுவனம் உற்பத்தி செய்யும் அனைத்து விளக்குகளையும் பற்றி முடிவு எடுப்பதுதான் சரியாக இருக்கும். எனவேதான், 'ஒரு பானை சோற்றுக்கு ஒரு சோறு பதம்' என்கிறார்கள்.

எடுக்கக்கூடிய மாதிரி (sample) அல்லது மாதிரிகள் சரியாக இருந்துவிட்டால், அவற்றின் மூலம் கிடைத்த, 'மாதிரி விபரங்கள்' (Statistics : characteristics of samples) மாதிரிகள் எடுக்கப்பட்ட முழுமையின் பண்பலகுகளை (parameters : characteristics of population) சரியாகக் காட்டும். எனவே, மாதிரி அல்லது மாதிரிகள் எடுப்பதற்கு சமவாய்ப்பு (random) முறையைப் பயன்படுத்த வேண்டும். சமவாய்ப்பு

முறையல்லாது வேறு முறைகளில் மாதிரிகள் எடுத்தால் அந்த ஆய்வு முடிவுகளை வைத்து பொதுமைப்படுத்துதல் சரியாக இருக்காது. பண்பலகுகளைச் சோதித்துப் பார்க்கவும் முடியாது; பண்பலகுகளை மதிப்பீடு செய்யவும் முடியாது. சமவாய்ப்பு முறையில் மாதிரி எடுப்பதற்கு முழுமையில் உள்ள அனைத்து உறுப்புக்களும் (elements) தெளிவாகவும் சரியாகவும் கிடைக்க வேண்டும். நிச்சயமற்ற உறுப்புக்களைக் கொண்ட முழுமையிலிருந்து சமவாய்ப்பு மாதிரி எடுப்பது சிரமம். எனவே, சமவாய்ப்பு மாதிரி எடுப்பதற்கு முதலில் முழுமையைச் சரியாக வரையறுத்துக் கொள்ள வேண்டும்.

தேவையான முன் ஏற்பாடுகள் எல்லாம் செய்துவிட்டு, சமவாய்ப்பு முறையில் மாதிரி எடுத்து, அந்த மாதிரியின் குணாதிசயங்களைக் (statistics) கணித்து அவற்றைப் பயன்படுத்தி முழுமையின் குணாதிசயங்களை (parameters) மதிப்பீடு (estimate) செய்வதன் மூலம் ஆய்வுகளின் பலனை அதிகரிக்க முடியும். எனவே, முதலில் மாதிரிகளைத் தேர்வு செய்வது பற்றிய கோட்பாடுகளைக் காணலாம்.

மாதிரிகள் எடுப்பது பற்றிய கோட்பாடு
(Sampling Theory)

சரியான மாதிரி எடுக்கும் முறைகளைப் (Methods of sampling) பற்றிய சோதனையை முதலில் திட்டமிட (design of the experiment) வேண்டும். சமவாய்ப்பு மாதிரி எடுப்பதற்கு, முதலில் முழுமையில் உள்ள கூறுகள் ஒவ்வொன்றுக்கும் தொடர்எண் கொடுத்து, குலுக்கல் முறையிலோ சமவாய்ப்பு எண்கள் முறையிலோ, கூறுகளைத் தெரிவு செய்ய வேண்டும். கூறுகளை எடுப்பதற்கு இரண்டு முறைகள் உள்ளன. ஒன்று எடுத்த கூறுகளை மீண்டும் முழுமையுடன் சேர்த்து (with replacement) மறுபடியும் மறுபடியும் கூறுகளை எடுப்பது; இதில், முதலில் எடுத்த கூறுகள் மீண்டும் வருவதற்கு வாய்ப்பு உண்டு. மற்றொன்று, எடுத்த கூறுகளை மீண்டும்

முழுமையுடன் சேர்க்காமல் அப்புறப்படுத்திவிட்டு (without replacement) வெவ்வேறு கூறுகள் வருவதற்கு வாய்ப்பளித்து எடுப்பது; இதில் எடுத்த கூறுகள் மீண்டும் வருவதற்கு வாய்ப்பு இல்லை.

முழுமை ஒரு முடிவுள்ளதாகவோ (finite) முடிவற்றதாகவோ (infinite) இருக்கலாம். உதாரணமாக, 100 பந்துகள் உள்ள ஒரு பெட்டியிலிருந்து, எடுத்த பந்துகளைத் திரும்பப் பெட்டிக்குள் போடாமல் 10 பந்துகள் கொண்ட மாதிரியாக மீண்டும் மீண்டும் மாதிரிகள் எடுத்தால், அது முடிவுள்ள முழுமையாகும். மாறாக, ஒரு நாணயத்தை 50 முறைகள் சுண்டிவிட்டு அதில் வரும் தலைகளைப் பற்றிய குறிப்பு எடுத்தால், அது முடிவற்ற முழுமையாகும்.

ஒரு முடிவுள்ள முழுமையிலிருந்து எடுத்த மாதிரிகளைத் திரும்பத் திரும்ப முழுமையுடன் சேர்த்து (with replacement) மாதிரி எடுத்தால், கோட்பாட்டுப்படி அது ஒரு முடிவற்ற முழுமையாகிறது. ஏனெனில் இந்த முறைப்படி அந்த முழுமையிலிருந்து எத்தனை மாதிரிகளை வேண்டுமானாலும் எடுக்கலாம். நடைமுறைக் காரணங்களுக்காக, மிகப்பெரிய முழுமைகள் முடிவற்ற முழுமைகளாகக் கருதப்படுகின்றன. எடுக்கப்படுகின்ற எல்லா மாதிரிகளுக்கும் (samples) அவற்றிற்குரிய தன்மைகளை (statistics) கண்டுபிடித்தால் ஒரு மாதிரிப் பரவல் (sampling distribution) கிடைக்கும். அது, பல மாதிரிகளிலிருந்து எடுக்கப்பட்ட சராசரிகளாக இருந்தால், மாதிரிகளின் சராசரிப் பரவல் (sampling distribution of means) என அழைக்கப்படுகிறது. இதுபோல மாதிரிகளின் திட்ட விலக்கங்களுக்கும், மாறுபாடுகளுக்கும் (variances) பரவல்கள் பெறலாம்.

மாதிரிகளின் சராசரிப் பரவல்
(Sampling distribution of means)

ஒரு முடிவுள்ள முழுமையிலிருந்து 'n' அளவுள்ள முடிந்த அளவுக்கு அனைத்து மாறிகளும் எடுக்கப்பட்டுள்ளன என வைத்துக்கொண்டால்,

$$\mu_{\bar{x}} = \mu \; ; \; \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$$

இதில், $\mu_{\bar{x}}$ = மாதிரிகளின் சராசரிப் பரவலின் சராசரி
(Mean of the sampling distribution of means)

μ = முழுமையின் சராசரி
(Parameter : Mean of the population)

$\sigma_{\bar{x}}$ = மாதிரிகளின் சராசரிப் பரவலின் திட்டவிலக்கம் (Standard deviation of the sampling distribution of means) இதற்கு 'திட்டப்பிழை' (standard error) என்ற பெயரும் உண்டு.

σ = முழுமையின் திட்டவிலக்கம்
(Standard deviation of the population)

n = மாதிரியின் அளவு (Sample Size)

N = முழுமையின் அளவு (Population size)

முடிவற்ற முழுமையாகவோ அல்லது எடுத்த மாதிரிகள் மீண்டும் மீண்டும் சேர்க்கப்பட்டு (with replacement) மாதிரிகள் எடுக்கப்பட்டு இருந்தாலோ,

$$\mu_{\bar{x}} = \mu \; ; \; \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \text{ ஆகும்.}$$

மாதிரியின் அளவு 30 ஆகவோ அல்லது 30க்கும் பெரியதாகவோ ($n \geq 30$) இருந்தால் மாதிரிகளின் சராசரிப் பரவல் ஓர் இயல்நிலைப் பரவலாக இருக்கும்; அந்தப் பரவல் $\mu_{\bar{x}}$யைச்

சராசரியாகவும் $\sigma_{\bar{x}}$-யைத் திட்ட விலக்கமாகவும் கொண்டிருக்கும். முழுமை எப்படியிருந்தாலும் மேலே கூறப்பட்டுள்ள உறவுகள் மாறாதிருக்கும்.

முழுமை ஓர் இயல்நிலைப் பரவலாக இருந்தால், மாதிரியின் அளவு 30யைவிடச் சிறியதாக இருந்தாலும், $(n < 30)$ மாதிரிகளின் சராசரிப் பரவலும் ஓர் இயல்நிலைப் பரவலாகவே இருக்கும்.

மாதிரிகளின் விகிதப் பரவல்
(Sampling distribution of proportions)

ஒரு முடிவில்லாத முழுமையில் இருவகையான நிகழ்வுகள் உள்ளதென வைத்துக் கொள்வோம். உதாரணமாக, ஒரு நாணயத்தை எண்ணற்ற முறையில் சுண்டுவதை எடுத்துக் கொள்ளலாம். அதில் தலை வருவதற்கான நிகழ்தகவும் (P) பூ வருவதற்கான நிகழ்தகவும் (Q) உள்ளன. அவை முறையே $\frac{1}{2}$யும் $\frac{1}{2}$யும் ஆகும். இவ்வாறாக இருக்கும்போது,

$$\mu_p = P; \ \sigma_p = \sqrt{\frac{PQ}{n}} = \sqrt{\frac{P(1-P)}{n}}$$ ஆகும்.

இதில் μ_p = மாதிரிகளின் விகிதப் பரவலின் சராசரி

σ_p = மாதிரிகளின் விகிதப் பரவலின் திட்டவிலக்கம்

n = மாதிரிகளின் அளவு

p = மாதிரிகளின் விகிதம் (தலை வருவதற்கான நிகழ்தகவு)

q = மாதிரிகளின் விகிதம் (பூ வருவதற்கான நிகழ்தகவு)

P = முழுமையில் தலை வருவதற்கான நிகழ்தகவு

Q = முழுமையில் பூ வருவதற்கான நிகழ்தகவு

இங்கு கொடுக்கப்பட்டுள்ள உதாரணம் ஈருறுப்புப் பரவலாக இருக்கும்போது, மாதிரியின் அளவு 30 ஆகவோ

30க்கும் பெரியதாகவோ ($n \geq 30$) இருந்தால், அது ஓர் இயல்நிலைப் பரவலை ஒத்திருக்கும்.

முடிவுள்ள முழுமையாக இருந்தாலும், மாதிரிகள் திரும்பச் சேர்க்கப்பட்டு (with replacement) எடுக்கப்பட்டால், மேலே கொடுக்கப்பட்டுள்ள உறவுகள் சரியாக இருக்கும்.

மாதிரிகளின் வித்தியாசங்கள் மற்றும் கூடுதல்களின் பரவல்கள் (Sampling distribution of differences and sums)

இரு வேறு முழுமைகளிலிருந்து, ஒவ்வொன்றிலும் ஒன்றாய், இரு சாரா மாதிரிகள் (independent samples) எடுக்கப்பட்டால் அவற்றிற்குள்ளே உள்ள உறவு கீழ்வருமாறு அமைந்திருக்கும்.

$$\mu_{\bar{x}_1-\bar{x}_2} = \mu_{\bar{x}_1} - \mu_{\bar{x}_2} = \mu_1 - \mu_2 \qquad \text{... (1)}$$

$$\sigma_{\bar{x}_1-\bar{x}_2} = \sqrt{\sigma^2_{\bar{x}_1} + \sigma^2_{\bar{x}_2}} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \qquad \text{... (2)}$$

$$\mu_{\bar{x}_1+\bar{x}_2} = \mu_{\bar{x}_1} + \mu_{\bar{x}_2} = \mu_1 + \mu_2 \qquad \text{... (3)}$$

$$\sigma_{\bar{x}_1+\bar{x}_2} = \sqrt{\sigma^2_{\bar{x}_1} + \sigma^2_{\bar{x}_2}} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \qquad \text{... (4)}$$

இரண்டு முழுமைகளிலிருந்து எடுக்கப்படும் சார்பிலா மாதிரிகளின் சராசரிகளுக்கிடையே உள்ள வித்தியாசங்கள் ஒரு பரவலாக இருக்கும். அந்தப் பரவலின் சராசரி, மாதிரிகள் எடுக்கப்பட்ட முழுமைகளின் சராசரிகளுக்கு இடையே உள்ள வித்தியாசத்திற்கு (difference) சமமாக இருக்கும் என்பதை முதல் சமன்பாடு *(1)* காட்டுகிறது. அதுபோல, இரண்டு முழுமைகளிலிருந்து எடுக்கப்படும் சார்பிலா மாதிரிகளின் சராசரிகளின் கூடுதல் ஒரு பரவலாக இருக்கும். அந்தப் பரவலின் சராசரி, மாதிரிகள் எடுக்கப்பட்ட முழுமைகளின்

சராசரிகளின் கூடுதலுக்கு சமமாக இருக்கும் என்பதை மூன்றாவது சமன்பாடு (3) காட்டுகிறது. அவற்றிற்கான திட்டப் பிழைகளை முறையே சமன்பாடு 2ம், சமன்பாடு 4ம் காட்டுகின்றன.

அட்டவணை - 45

சில மாதிரிகளின் பரவல்களுக்கான திட்டப்பிழைகள்

மாதிரிகளின் பரவல் (Sampling distribution)	திட்டப்பிழை (Standard error)	குறிப்பு (Remarks)
கூட்டுச் சராசரிகள் (Means)	$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$	$n \geq 30$ n = மாதிரியின் அளவு
விகிதங்கள் (Proportions)	$\sigma_p = \sqrt{\frac{PQ}{n}}$	$n \geq 30$ n = மாதிரியின் அளவு
திட்டவிலக்கங்கள் (Standard deviations)	$\sigma_s = \frac{\sigma}{\sqrt{2n}}$	$n \geq 100$ n = மாதிரியின் அளவு
மாறுபாடுகள் (Variances)	$\sigma_s^2 = \sigma^2\sqrt{\frac{2}{n}}$	$n \geq 100$ n = மாதிரியின் அளவு

முழுமையின் அலகுகளை மதிப்பீடு செய்தல் (Estimation of Parameters)

இதுவரையில் மாதிரிகளின் குணங்களுக்கும் (Statistics) முழுமையின் குணங்களுக்கும் இடையே உள்ள தொடர்புகள் விவரிக்கப்பட்டன. ஒரு முழுமையின் குணங்கள் (parameters) தெரியவில்லையென்றால், மேலே கூறிய விபரங்களைக் கொண்டு, ஒரு முழுமையின் குணங்களையோ அல்லது இரண்டு அல்லது இரண்டுக்கும் மேற்பட்ட முழுமைகளின்

குணங்களுக்கிடையேயுள்ள உறவுகளையோ மதிப்பீடு செய்யலாம். இவ்வாறு செய்யும் செயலுக்கு புள்ளியியல் தீர்மானங்கள் அல்லது அநுமானங்கள் (Statistical inferences) என்று பெயர்.

மாதிரியிலிருந்து பெறப்பட்ட அலகுகள் (Statistics) முழுமையின் அலகுகளுக்குச் (Parameters) சமமாக இருந்தால், அவை பேதமற்ற (Unbiased) மதிப்பீடுகள் (estimators) என அழைக்கப்படுகின்றன. மாதிரி அலகுகளிலிருந்து பெறப்பட்ட முழுமையின் அலகுகள் மதிப்புக்கள் (estimates) என அழைக்கப்படுகின்றன.

உதாரணமாக, மாதிரியின் கூட்டுச் சராசரி ($\overline{X}$), அந்த மாதிரி எடுக்கப்பட்ட முழுமையின் கூட்டுச் சராசரியுடைய (μ) பேதமற்ற மதிப்பீடு ஆகும்.

அதுபோல மாதிரிகளின் மாறுபாட்டுப் பரவலின் கூட்டுச் சராசரி (Mean of the sampling distribution of variances) $\mu_s{}^2 = \frac{(n-1)\sigma^2}{n}$. இதில் σ^2 என்பது முழுமையின் மாறுபாடு (variance of population); n என்பது மாதிரியின் அளவு (sample size). இவ்வாறாக, s^2 என்பது σ^2ன் பேதமுள்ள மதிப்பீடு (biased estimate) ஆகும். எனவே இந்த s^2யைச் சிறிது மாற்ற வேண்டும் $s^2 = \frac{\Sigma (X-\overline{X})^2}{n}$ என்று பெறப்படுகிறது. இது பேதமாக உள்ளது. இதில் nக்குப் பதிலாக n-1யைக் கொண்டு வகுத்தால் பேதம் நீங்கிவிடும். எனவே s^2 முதலிலேயே கண்டுபிடிக்கப்பட்டிருந்தால் அதை nஆல் பெருக்கி (ns^2) n-1ஆல் வகுத்து விடவேண்டும். இப்படி மாற்றம் செய்த பின்னர் கிடைக்கும் s_2யை $\hat{s}^2$ என்று அழைக்கிறார்கள். இந்த $\hat{s}^2$ ஒரு பேதமற்ற மதிப்பீடாகும். எனவே $\mu_{\hat{s}^2} = \sigma^2$. ஆனாலும் $\hat{s}$என்பது σன் ஒரு பேதமுள்ள மதிப்பீடாகவே இருக்கிறது.

இவ்வாறாக $\overline{X}$ என்பது μன் ஒரு பேதமற்ற மதிப்பீடாகவும், $\hat{s}^2$ என்பது σ^2ன் பேதமற்ற மதிப்பீடாகவும் இருக்கின்றன. இவற்றைச் சுருக்கமாக $E(\overline{X}) = \mu$ என்றும், $E(\hat{s}^2) = \sigma^2$ என்றும் சொல்லலாம்.

திறன்மிக்க மதிப்பீடு

மாதிரிகளிலிருந்து கிடைத்த இரண்டு அலகுகளுமே (Statistics) சமமான கூட்டுச் சராசரியைக் கொண்டிருந்தால், எந்த அலகு குறைவான மாறுபாட்டை (variance) கொண்டுள்ளதோ, அந்த அலகு திறன்மிக்க மதிப்பீடாகக் (efficient estimator) கருதப்படுகிறது. அவற்றிற்குத் தொடர்பான மதிப்புக்கள் (estimates) திறன் குறைந்த மதிப்புக்கள் (inefficient estimates) என அழைக்கப்படுகின்றன. மிகமிகக் குறைவான மாறுபாட்டைக் கொண்டிருக்கும் மதிப்பீடு மிகச்சிறந்த (Most efficient or best) மதிப்பீடாக கருதப்படுகிறது.

மதிப்புப் புள்ளியும் மதிப்பு இடைவெளியும் (Point estimates and Interval estimates)

மாதிரியிலிருந்து கிடைக்கின்ற அலகுகளைக் (statistics) கொண்டு, முழுமையின் பண்பலகினை (parameters) மதிப்பீடு செய்யும்போது, சரியாக ஒரே ஒரு எண்ணை மட்டும் (உ.ம்.100) கொடுத்தால் அது மதிப்புப்புள்ளி என்று அழைக்கப்படுகிறது. அவ்வாறு கொடுக்கும்போது, இரண்டு எண்களைக் கொடுத்து (உதாரணமாக 100 முதல் 150 வரை) இவ்விரண்டு எண்களுக்கிடையில்தான் முழுமையின் பண்பலகு இருக்கும் என மதிப்பீடு செய்தால் அதற்கு மதிப்பு இடைவெளி என்று பெயர்.

நம்பிக்கை இடைவெளி (Confidence intervel)

ஒரு நிறுவனத்தின் பங்கின் (Share) மதிப்பு ஒரு நாள் ரூ.100 இருக்கின்றதென்று கொள்வோம். அதற்கு அடுத்த நாள் அந்தப் பங்கின் மதிப்பு எவ்வளவு இருக்கும் என்பதை 100

விழுக்காடு துல்லியமாகச் சொல்வது என்பது மிகவும் சிரமம். அடுத்த நாள் அந்தப் பங்கின் விலை ரூ.105 ஆக இருக்கும் என்று சொல்லும்போது அவ்வாறு சொல்பவரின் மனதில் உள்ள நம்பிக்கையளவு மிகக் குறைவாக இருக்கும். மற்றொருவர் அந்தப் பங்கின் விலை ரூ.100 முதல் ரூ.110க்குள் இருக்கும் என்று சொன்னால் அவரின் நம்பிக்கையளவு முன்னவரின் நம்பிக்கையளவை விடச் சற்று அதிகமாக இருக்கும். இன்னொருவர் அந்தப் பங்கின் விலை ரூ.90 முதல் ரூ.120 வரை இருக்கும் என்று சொல்லும்போது அவரின் நம்பிக்கையளவு இன்னும் அதிகமாக இருக்கும்.

அட்டவணை - 46

நம்பிக்கையின் அளவு சதவீதத்தில் (%)	99	98	95	90	50
நிலைப்படுத்தப்பட்ட அளவு z_c	2.58	2.33	1.96	1.645	0.67

அதையே, புள்ளியியல் முறைப்படிக் கூறினால், முழுமையின் சராசரி $\overline{X} \pm 1.96$ SE (SE = Standard Error)க்குள் இருக்கும் எனும்போது இடைவெளி மிகக் குறுகலாக உள்ளது. இதையே $\overline{X} \pm 2.58$ SEக்குள் இருக்கும் என்று சொல்லும்போது இடைவெளி அதிகமாகிறது; எனவே நம்பிக்கையின் அளவும் அதிகமாகிறது. இதில் முன்னது 95 சதவீத நம்பிக்கை அளவு என்றும் பின்னது 99 சதவீத நம்பிக்கை அளவு என்றும் கூறப்படுகின்றன. 95%, 99% ஆகியவை நம்பிக்கை அளவுகள் (Confidence levels) எனவும், $\overline{X} \pm 2.58$ SE, $\overline{X} \pm 2.58$ ஆகியவை நம்பிக்கை எல்லைகள் அல்லது வரம்புகள் (Confidence limits or fiducial limits) எனவும் அழைக்கப்படுகின்றன. 1.96, 2.58 ஆகியவை நம்பிக்கைக் கெழுக்கள் (Confidence coefficients or critical values) எனவும் அழைக்கப்படுகின்றன. பொதுவாக $\overline{X} \pm z_c s_{\bar{x}}$ என்று சராசரிக்கும், $p \pm z_c \sqrt{\frac{PQ}{n}}$ என்று விகிதத்திற்கும்

நம்பிக்கை எல்லைகள் கொடுக்கப்படுகின்றன. இரண்டு முழுமைகளின் சராசரிகளுக்குள்ளே உள்ள வித்தியாசங்களுக்கு

$$\bar{X}_1 - \bar{X}_2 \pm z_c SE$$ என்றும்

இரண்டு முழுமைகளின் சராசரிகளின் கூடுதலுக்கு

$$\bar{X}_1 + \bar{X}_2 \pm z_c SE$$ என்றும்

இரண்டு முழுமைகளின் விகிதங்களுக்கு இடையே உள்ள வித்தியாசங்களுக்கு

$$p_1 - p_2 \pm z_c SE$$ என்றும்

இரண்டு முழுமைகளின் விகிதங்களின் கூடுதலுக்கு

$$p_1 + p_2 \pm z_c SE$$ என்றும்

ஒரு முழுமையின் திட்டவிலக்கத்திற்கு

$$s \pm z_c SE$$ என்றும்

நம்பிக்கை எல்லைகள் கொடுக்கப்படுகின்றன.

எடுகோள்களைச் சோதித்தல் (Tests of Hypotheses)

ஒன்று அல்லது ஒன்றுக்கு மேற்பட்ட கேள்விகள் எழும்போது அவற்றிற்கான பதிலைத் தேடுவதற்கு ஓர் ஆய்வினை மேற்கொள்கிறார்கள். எழுப்பப்பட்ட அல்லது எழுந்த வினாக்களுக்கான பதிலுக்காக பல செய்திகள் சேகரிக்க வேண்டியுள்ளது. அந்த செய்திகளைச் சேகரிக்கச் செல்லும்போது சில மாறிகளைப் பற்றிய செய்திகளும் அந்த மாறிகளுக்கு இடையேயான உறவுகளும், தொடர்புகளும் தெரிய வருகின்றன. இச்செய்திகளும் மாறிகளுக்கிடையேயான உறவுகளும், முன்னரே செய்த ஆய்வுகளிலிருந்தும் தெரிய வரலாம். அன்றாட வாழ்வு முறையில், கண்ணுக்கெதிரே நடக்கும் நிகழ்ச்சிகளிலிருந்தும் தெரிய வரலாம். கிடைத்த செய்திகள், புள்ளி விபரங்கள், அவற்றிற்கு இடையேயுள்ள

உறவு முறைகளை ஆழ்ந்து சிந்திக்கும்போது, சில சமயம் முன்னர் ஆய்வின் மூலம் கூறப்பட்ட கருத்துக்கள் இன்னமும் சரியாகத் தோன்றலாம்; அல்லது பல துறைகளிலும் ஏற்பட்ட மாற்றங்களினால் முன்னர் கூறப்பட்ட கருத்துக்களில் சில மாற்றங்கள் செய்யப்பட வேண்டிய தேவைகள் எழலாம்; அல்லது, முன்னர் ஏதோ ஒரு குறிப்பிட்ட சூழலுக்குப் பொருத்தமாக இருந்த கருத்து வேறொரு சூழலில் சரியில்லாமல் இருக்கலாம். இவைகளாலும் இன்னும் பல காரணங்களினாலும், மாறிகளுக்கிடையேயான உறவுகளும் மாறிக்கொண்டேயிருக்கும். இத்தருணத்தில், 'மாறும் என்ற உண்மையைத் தவிர மற்றவையெல்லாம் மாறிக் கொண்டிருக்கும்' என்ற பொருத்தமான உண்மையை நினைத்துப் பார்க்கலாம். இவ்வாறு மாற்றங்கள் தொடர்ந்து நிகழ்ந்து கொண்டே இருப்பதாலும், எல்லா மாற்றங்களும் எல்லாருக்கும் சமமான சாதக பாதக விளைவுகளைக் கொண்டுவர இயலாததாலும், தொடர்ந்து ஆய்வுகள் நடந்து வருகின்றன. எவ்வளவுதான் புதியவை புனைந்தாலும், அவை பழையன ஆதலாலும்; புதியவை தொடர்ந்து புகுந்து கொண்டே இருப்பதாலும், புதியன புகுதலும் பழையன கழிதலும் இயற்கையின் நியதி என்பார்கள். எனவே, ஆய்வுக்கு முழுமையும் வராது; புதுமையும் முடியாது; தேவை தோன்றிக்கொண்டே இருக்கும்.

எடுகோளுக்கும் (Hypothesis) அநுமானத்திற்கும் (Assumption) வேறுபாடுகள் உள்ளன; இரண்டும் ஒன்றல்ல. சுருக்கமாகச் சொல்லப்போனால் சில அநுமானங்களின் அடிப்படையில் எடுகோள்கள் உருவாக்கப்படுகின்றன. உதாரணமாக ஒரு பொருளின் தேவைக்கும் அதன் விலைக்கும் எதிரிடையான உறவு உள்ளது என்பது ஒரு எடுகோள். இந்த எடுகோள், மற்றவைகள் (மற்ற பொருள்களின் விலைகள், நுகர்வோரின் வருமானம், வாழ்வுமுறை ... போன்றவைகள்)

மாறவில்லை என்னும் அநுமானத்தின் அடிப்படையில் உருவாக்கப்படுகிறது.

இன்னும் ஓர் உதாரணம் எடுகோளுக்குக் கொடுக்க வேண்டும் எனில், மாணவர்களின் படிப்புத்திறமைக்கும் மாணவிகளின் படிப்புத்திறமைக்கும் வேறுபாடு இல்லை எனும் இல்லெனும் எடுகோளைக் கூறலாம். மாணவ மாணவிகளின் படிப்புத்திறமையை ஒப்பிட அவர்கள் பெற்ற மதிப்பெண்களைப் பயன்படுத்தலாம். இங்கும் சில அநுமானங்கள் அடிப்படையாக இருக்கிறன. அவை:

1. மதிப்பெண்கள் படிப்புத்திறமையை காட்டமுடியும்.
2. மாணவர்களுக்கும் மாணவிகளுக்கும் பாரபட்சமின்றி மதிப்பெண்கள் வழங்கப்பட்டுள்ளன.
3. மாணவர்களுக்கும் மாணவிகளுக்கும் ஒரேவிதமான கேள்விகளே ஒரேவிதமான அமைப்பிலேயே அளிக்கப்பட்டுள்ளன.
4. மாணவர்களுக்கும் மாணவிகளுக்கும் தேர்வு நடந்த அறையும் தேர்வு நடந்த நேரமும் ஒரே மாதிரி இருந்தன.
5. மாணவர்களுக்கும் மாணவிகளுக்கும் பாடம் நடத்திய ஆசிரியர்கள் எல்லா வகையிலும் சமமான தகுதி பெற்றவர்களே.
6. மாணவர்களுக்கும் மாணவிகளுக்கும் படிப்புச் சூழ்நிலையும், பாடப்புத்தகங்களும், குடும்பச் சூழலும் சம அளவாகச் சம தன்மையாக இருந்தன.

இவை போன்ற இன்னும் பல அநுமானங்களின் அடிப்படையில் தான் மேலே கூறப்பட்டுள்ள இல்லெனும் எடுகோள் உருவாக்கப்பட்டுள்ளது. எனவே எடுகோள் வேறு; அநுமானம் வேறு.

இல்லெனும் எடுகோள் (Null Hyothesis)

பொதுவாக நீண்ட நாட்களாக வாழ்நாளில் ஆய்வு செய்துள்ள குன்னார் மிர்தால் (GUNNAR MYDRAL) போன்றவர்கள் முன்னரே சேகரிக்கப்பட்ட விழுமியங்கள், மதிப்பீடுகள் (values) போன்றவற்றில் மூழ்கி விடாமல் (pre - conceived notions) கவனமாக எந்தவித பேதமும் (bias) இல்லாமல் நேர்மை பிறழாது (objective) ஆய்வினைத் தொடங்க வேண்டும் என்பார்கள். முன்னரே ஒரு கருத்துக்கு அடிமையாகி விட்டால், அதைப் பற்றிய உண்மையான செய்திகளைச் (true information) சேகரிப்பதற்குப் பதிலாக தமக்குத் தேவைப்படும் நல்ல செய்திகளை (good imformation) மட்டுமே சேகரிக்கத் தொடங்கிவிட வாய்ப்புண்டு என்கிறார்கள். எனவேதான், ஆய்வினை இல்லெனும் எடுகோளுடன் (Null Hypothesis) தொடங்க வேண்டும் என்கிறார்கள். இதற்கு உதாரணமாக, 'மாணவர்களுக்கும் மாணவிகளுக்கும் உயரத்தில் வித்தியாசம் இல்லை' என்பதை இல்லெனும் எடுகோள் எனலாம். படிப்பறிவுக்கும் நல்ல குணங்களுக்கும் தொடர்பில்லை என்பதும் இல்லெனும் எடுகோளுக்கு உதாரணமாகும். இவ்வாறாக வித்தியாசம், வேறுபாடு, தொடர்பு, உறவு இல்லை என்பதைக் கொண்டு இல்லெனும் எடுகோள்கள் உருவாக்கலாம். வேறுவிதமாகவும் வார்த்தைகளைப் பயன்படுத்தி எடுகோள்கள் உருவாக்கலாம். உதாரணமாக, இரண்டு மாறிகளிடையே உறவில்லை என்பதற்குப் பதிலாக, அவை இரண்டும் சார்பிலாதவை (independent) என்றும் சொல்லலாம். இரண்டு மாறிகளுக்கு இடையில் வேறுபாடு இல்லை என்பதை இரண்டு மாறிகளும் சமமாக இருக்கின்றன என்றும் சொல்லி இல்லெனும் எடுகோள்கள் உருவாக்கலாம்.

மாற்று எடுகோள் (Alternative Hypothesis)

கிடைக்கக்கூடிய செய்திகள் மற்றும் புள்ளி விபரங்கள் மூலம் இல்லெனும் எடுகோளை ஏற்றுக் கொள்ளவும்

(accepting null hypothesis) செய்யலாம்; ஏற்றுக் கொள்ளாமலும் (not accepting) இருக்கலாம். இல்லெனும் எடுகோளை ஏற்றுக்கொள்ளவில்லையெனில், அதன் மாற்றான மாற்று எடுகோளை (alternative hypothesis) ஏற்றுக் கொள்ளலாம் எனப்படும். அப்படியானால், மாற்று எடுகோள்களை எப்படி அமைப்பது என்பது பற்றி இனிக் காணலாம். மாற்று எடுகோள்களை மூன்று விதமாக அமைக்கலாம். மிகச் சுருக்கமாக வித்தியாசங்களைப் பொறுத்தமட்டில்,

$H_1 : P_1 \neq P_2$; $\mu_1 \neq \mu_2$; $P \neq 90$; $\mu \neq 5.6$ ... (1)

$H_1 : P_1 > P_2$; $\mu_1 > \mu_2$; $P > 90$; $\mu > 5.6$ ... (2)

$H_1 : P_1 < P_2$; $\mu_1 < \mu_2$; $P < 90$; $\mu < 5.6$ ... (3)

என்று சொல்லலாம்.

உறவுகளைப் பொறுத்தமட்டில்,

H_1 : இரண்டுக்கும் தொடர்பு உள்ளது

H_1 : இரண்டும் ஒன்றை ஒன்று சார்ந்துள்ளன

H_1 : இரண்டும் சாரா நிகழ்வுகள் அல்ல
(Dependent or not independent)

மேலேகூறப்பட்டுள்ள மாற்று எடுகோள்களைப் பொறுத்து, ஒருபுறச் சோதனையா (single tail test) இருபுறச் சோதனையா (two tail test) என்று முடிவு செய்ய வேண்டிவரும். ஒருபுறச் சோதனை என்றால் இடதுபுறச் (left tail) சோதனையா வலதுபுறச் சோதனையா (right tail) என்று முடிவு செய்ய வேண்டியிருக்கும். முதலில் கொடுக்கப்பட்டுள்ள மூன்று வகையான மாற்று எடுகோள்களில் முதலாவது வகைக்கு இருபுறச் சோதனை பொருத்தமாகும். இரண்டாவது வகைக்கு வலதுபுறச் சோதனை பொருத்தமாகும். மூன்றாவது வகைக்கு இடதுபுறச் சோதனை பொருத்தமாகும்.

அடுத்ததாக எப்படிப்பட்ட சூழ்நிலைகளில் இருபுறச் சோதனை வேண்டும், எப்படிப்பட்ட சூழ்நிலைகளில் இட/வலப்புறச் சோதனைகள் வேண்டும் என்று தெரிந்து கொள்வது அவசியம். உதாரணமாக 100 நாணயங்களைக் குலுக்கிப் போட்டால் ஏதேனும் 50 நாணயங்களில் தலை மேலே வரவேண்டும் என்று எதிர்பார்க்கிறோம். அப்படியானால், மீதமுள்ள 50 நாணயங்களில் பூ மேலே வரவேண்டுமெனப் பொருள். அதுபோல ஒரு நாணயத்தை நூறு தடவைகள் சுண்டிவிட்டால் ஏதேனும் 50 தடவைகள் தலை வந்தால் அதனை நல்ல பேதமற்ற நாணயம் எனலாம். அப்படியல்லாமல், 30 தடவைகள் தலை வந்தாலும், 70 தடவைகள் தலை வந்தாலும் அந்த நாணயம் நல்லதல்ல என்றே முடிவு செய்வோம். எனவே, தலைகளின் எண்ணிக்கை 50க்கு மேலிருந்தாலும் சரி, கீழிருந்தாலும் சரி, முடிவு ஒன்றுதான்; அதாவது நாணயம் நல்லதல்ல. இப்படிப்பட்ட சூழ்நிலையில் மாற்று எடுகோள் (1) பொருத்தமாகும். அதுபோல, விமானப் பணிப் பெண்ணுக்குத் தேவையான உயரம் ஐந்து அடி ஆறு அங்குலம் என்று நிர்ணயிக்கப்பட்டால் ஐந்து அடி ஒன்பது அங்குலம் உள்ள பெண்ணும் நிராகரிக்கப்படலாம்; ஐந்து அடி 3 அங்குலம் உள்ள பெண்ணும் நிராகரிக்கப்படலாம். இந்தச் சூழ்நிலைகளிலும் முதலாவது வகை எடுகோள் பொருந்தும். தேசிய மாணவியர் படைத் தேர்வுக்குத் தேவையான உயரம் ஐந்து அடி ஆறு அங்குலம் என்று நிர்ணயிக்கப்பட்டால் குறிப்பிடப்பட்டுள்ள உயரத்தைவிட அதிக உயரம் உள்ள மாணவி தேர்ந்தெடுக்கப்படுவார்; அதைவிடக் குறைவான உயரம் உள்ள மாணவி மட்டும்தான் நிராகரிக்கப்படுவார். இந்த சூழ்நிலையில் இரண்டாவது வகையான மாற்று எடுகோள் பொருந்தும்.

ஒருவர் தான் விற்கும் மாம்பழக் கூடையில் உள்ள 100 மாம்பழங்களில் அதிகமாகப் போனால் 10 பழங்கள்தான்

அழுகியிருக்கலாம் என்று சொல்கிறார். ஆனால் கூடையைத் திறந்து பார்த்தபோது 5 மாம்பழங்கள் தான் அழுகி உள்ளன. இருந்தாலும் அவர் வாக்கு சரியென்போம்; ஆனால், அழுகிய பழங்கள் 10யை விடக் கூடுதலாக (உ.ம்.20) இருந்தால் அவருடைய வாக்கு தவறிவிட்டது என்போம். இதில் கூடுதலாக இருந்தால் நிராகரிக்கிறோம். எனவே, இதற்கு மூன்றாவது வகையான மாற்று எடுகோள்தான் பொருந்தும். ஆனாலும், இதற்கு இரண்டாவது வகையான மாற்று எடுகோள் பொருந்துவது போல் வார்த்தைகளை மாற்றிப் போடலாம். உதாரணமாக, மாம்பழம் விற்பவர் தன் கூடையில் உள்ள பழங்களில் 90 பழங்கள் நல்ல பழங்கள் என்று வாக்குக் கொடுத்திருந்து, அதில் 95 பழங்கள் நல்ல பழங்களாக இருந்தால், அவர் கூறிய வாக்குறுதி சரியாகி விடும். இப்படிப்பட்ட சூழ்நிலையில் இரண்டாவது வகையான மாற்று எடுகோள் அமைப்பது சரியாகும்.

பிழைகள்

'பிழைத்தல் இன்றி பிழைப்பு இல்லை' என்பார்கள். என்ன செயல் செய்தாலும் பிழைப்பதற்கு (to err) வாய்ப்பு உள்ளது. ஒவ்வொரு முடிவும் சரியாக இருப்பதற்கும் வாய்ப்பு உள்ளது; தவறாக முடிவதற்கும் வாய்ப்பு உள்ளது. உதாரணமாக, ஒருவர் ஒரு குலுக்குச் சீட்டு மூலம் பணம் சம்பாதிக்க ஆசைப்பட்டு குலுக்குச் சீட்டு ஒன்று வாங்குகிறார் என்றால் அது தவறாகவும் முடியலாம்; நல்லதாகவும் முடியலாம். அதுபோல், அவர் அந்த குலுக்குச் சீட்டை வாங்காமல் போவது கூட நல்லதாகவும் இருக்கலாம் (ஏனெனில் பரிசு கிடைக்கவில்லையெனில் நட்டத்தைத் தவிர்க்கலாம்); அந்தச் சீட்டை வாங்காமல் போனது பிழையாகக் கூட இருக்கலாம். (ஏனெனில், சில நாட்கள் கழித்துப் பார்த்தால் அவர் வாங்காமல் விட்டுச் சென்ற சீட்டுக்கு பரிசு கிடைத்திருக்கலாம்). இதுபோல, பல உதாரணங்கள் கூற முடியும்.

வீட்டைவிட்டு வெளியே செல்லும்போது குடை எடுத்துக்கொண்டு போனால் (மழை பெய்தால்) நல்லதாகவும் இருக்கலாம்; தவறாகவும் (மழை பெய்யவில்லையென்றால் குடையை எங்காவது மறந்து வைத்துவிட்டு வருவதற்கு அதிக வாய்ப்பு உள்ளதால்) இருக்கலாம்.

ஒரு சாலையில் செல்லும்போது வண்டியை நிறுத்துவதற்கான சிவப்பு விளக்கு தெரியும்போது நிற்காமல் வண்டியைத் தொடர்ந்து ஓட்டிக் கொண்டு போவது தவறாகி விடலாம்; அல்லது அங்கு பச்சை விளக்கு தெரியும்போது, வண்டியை நிறுத்தியும் வைக்கலாம். இந்த இரண்டுமே பிழைகள்தான். இந்தப் பிழைகளுள் எந்தப் பிழையினால் அதிக நட்டம் / சிரமம் வரும் எனப் பார்க்க வேண்டும்.

மிக நேர்த்தியான வணிகர் நல்ல பொருளை மட்டுமே சந்தைக்கு அனுப்ப வேண்டும் என நினைப்பார். அவர் சில நல்ல பொருள்கள் அங்காடிக்குச் செல்லாமல் தடுக்கப்பட்டாலும் பரவாயில்லை; ஒரு பழுதான பொருள் கூட சந்தைக்குச் செல்லக்கூடாது என்பதில் கண்ணும் கருத்துமாக இருப்பார். ஆனால் விரைவில் பணம் சம்பாதிக்க நினைக்கும் வணிகர் பல பழுதான பொருள்கள் சந்தைக்குப் போனாலும் பரவாயில்லை; ஒரு பொருள் கூட சந்தைக்குப் போகாமல் இருக்கக்கூடாது என நினைக்கலாம். பழுதான பொருளைச் சந்தைக்கு அனுப்புவதும் தவறுதான் (ஏனெனில் அவரின் பொருள்கள் மேல் அவநம்பிக்கை வந்து விற்பனை குறைந்து நட்டம் வர வாய்ப்புள்ளது). நல்ல பொருள்களைச் சந்தைக்கு அனுப்பாமல் (நம்பிக்கை குறைந்துவிடக் கூடாது என்பதற்காக) வைப்பதும் நட்டத்தை விளைவிக்கலாம். இரண்டுமே பிழைகள் என்றாலும் எந்தப்பிழை செய்வது சரி என்று முடிவு செய்வது அவரவர் மனநிலையைப் பொறுத்தது. நல்ல பொருளைச் சந்தைக்கு அனுப்பாததும் பிழை. பழுதான பொருளைச் சந்தைக்கு அனுப்புவதும் பிழை. இதில் ஒன்றை முதல்விதப் பிழை (TYPE I ERROR) எனவும்,

மற்றொன்றை இரண்டாம்விதப் பிழை (TYPE II ERROR) எனவும் சொல்லலாம்.

நல்லதை மறுப்பது முதல்விதப் பிழை (Rejecting the true - Type I) தவறானதை ஒத்துக் கொள்வது இரண்டாம்விதப் பிழை (Accepting the false - Type - II)

இரண்டு வகையான பிழைகளை திருவள்ளுவர்,

"செய்தக்க வல்ல செயக்கெடுஞ் செய்தக்க
செய்யாமை யானுங் கெடும்"

என்று தெரிந்து செயல்வகை அதிகாரத்தில் கூறியுள்ளார்.

இந்த இருவகையான பிழைகளையும் எடுகோள்களை வைத்தும் கூறலாம். ஒன்று உண்மையான (True) எடுகோள் என்றும் மற்றொன்று தவறான (False) எடுகோள் என்றும் கொள்வோம். உண்மையான எடுகோளை மறுப்பது முதல்வகைப் பிழை (Type I Error) தவறான எடுகோளை சரியென ஒத்துக் கொள்வது இரண்டாவது வகைப் பிழை (Type II Error). இவ்விரண்டில், முதல்வகைப் பிழையைக் குறைக்க முயற்சிக்கும்போது இரண்டாம்வகைப் பிழையைச் செய்ய வாய்ப்புக் கூடும். அதுபோலவே இரண்டாம்வகைப்பிழையைக் குறைக்க முயற்சிக்கும்போது முதல் வகைப் பிழையைச் செய்ய வாய்ப்புக் கூடும். இரண்டு பிழைகளையும் குறைக்க ஒரே வழி மாதிரியின் அளவைக் (sample size) கூட்டுவதுதான்.

குற்றவாளிகளை நிரபராதி என்று விடுதலை செய்வதும் பிழை; நிரபராதிகளைக் குற்றவாளிகள் என்று தண்டிப்பதும் பிழை. ஆனாலும், ஆயிரம் குற்றவாளிகள் தப்பித்துச் சென்றாலும் பரவாயில்லை; ஒரு நிரபராதி கூட தண்டிக்கப்படக் கூடாது என்பது நீதிமுறை என்பார்கள்.

எனவே, உண்மையான (True) எடுகோளை ஒத்துக் கொள்வதற்கான வாய்ப்பினை அதிகமாக்கி, அந்த உண்மையான எடுகோளை மறுப்பதற்கான வாய்ப்பினை குறைக்கவே முயற்சிக்கிறோம்.

முக்கியத்துவ நிலைகள் (Levels of Significance)

முக்கியத்துவ நிலைகளாக, காலம்காலமாக, 10 சதவீதம், 5 சதவீதம், 1 சதவீதம் என்று, வசதிக்காக எடுத்துக் கொள்ளப்படுகின்றன. இவற்றை நன்றாகப் புரிந்துகொண்டால், மற்ற சதவீதங்களைப் (உதாரணமாக : 8%, 2%, 0.5%) பற்றி எளிதாகப் பிறகு புரிந்து கொள்ள முடியும். முக்கியத்துவ நிலையாக 10 சதவீதத்தை எடுத்துக் கொண்டால், 10 சதவீதம் அளவுக்குப் பிழை செய்கிறோம் என்று பொருள். அதாவது, எடுக்கப்படும் முடிவு 90 சதவீதம் சரியாக இருக்கும் என்றும் 10 சதவீதம் தவறாகப் போகலாம் என்றும் பொருள். இன்னொரு முறையில், 90 சதவீதம் எடுக்கப்பட்ட முடிவு சரியாக இருக்கும் என்று நம்பலாம். 90 சதவீத நம்பிக்கை அளவு என்றால், 10 சதவீத முக்கியத்துவ நிலையென்று பொருள். அதாவது, எடுக்கப்படும் முடிவு தவறாக இருப்பதற்கு நிகழ்தகவு 0.10 என்று அர்த்தம் ஆகும்.

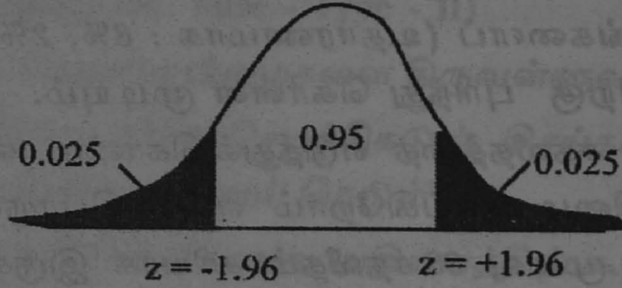
அட்டவணை - 47

முக்கியத்துவநிலை	10	8	5	4	2	1	0.50
நம்பிக்கை நிலை (%)	90	92	95	96	98	99	99.50

அதுபோல, முக்கியத்துவ நிலை 5 சதவீதம் என்றால், எடுத்த முடிவு தவறாக இருப்பதற்கு வாய்ப்பு 5 சதவீதம் என்று பொருள்; அல்லது, எடுத்த முடிவு தவறாக இருப்பதற்கான நிகழ்தகவு 0.05 என்று பொருள். இன்னொரு வகையில் சொன்னால், எடுத்த முடிவு சரியாக இருப்பதற்கு 95 சதவீத வாய்ப்புக்கள் உள்ளன. அல்லது எடுத்த முடிவு சரியாக இருப்பதற்கான நிகழ்தகவு 0.95 ஆகும்.

இயல்நிலைப் பரவல் கொண்ட சோதனைகள் (Tests involving the normal distribution)

வரைபடம் - 38



வரைபடம் 38இல் காட்டப்பட்டதுபோல், 95 சதவீதப் பரப்பளவு z = ± 1.96க்குள் உள்ளது. மீதமுள்ள 5 சதவீதப் பரப்பு இடதுபுறமும் வலதுபுறம் 2.5 சதவீதமாக பரவியுள்ளது. இதிலிருந்து, நாம் எடுக்கும் கணக்கில் கண்டுபிடிக்கப்படும் நிலைப்படுத்தப்பட்ட மதிப்பு (z score) ± 1.96க்குள் வந்தால், நாம் எடுத்துள்ள இல்லெனும் எடுகோளை (null hypothesis) ஒத்துக் கொள்ளலாம். நிலைப்படுத்தப்பட்ட மதிப்பு -1.96க்கும் கீழ் அல்லது +1.96க்கும் மேல் வந்தால், இல்லெனும் எடுகோள் நிராகரிக்கப்படும்.

உதாரணமாக, நாம் இரண்டு கூட்டுச் சராசரிகளை ஒப்பிட்டு இரண்டுக்குமிடையே வித்தியாசம் உள்ளதா? அப்படி வித்தியாசம் இருந்தால், அது புள்ளியியல் முக்கியத்துவம் (statistical significance) பெறுகிறதா? என்று பார்க்க விரும்புகிறோம் என வைத்துக் கொள்வோம். இதற்கு, இல்லெனும் எடுகோள் (null hypothesis) $H_0 : \mu_1 = \mu_2$ என்று எடுக்கலாம். மாற்று எடுகோள் (alternative hypothesis) $H_1 : \mu_1 \neq \mu_2$ என்று எடுக்கலாம். கிடைத்திருக்கும் புள்ளி விபரங்களைக் கொண்டு zன் மதிப்பு கண்டுபிடிக்கலாம். அந்த zன் மதிப்பு 3.2 என வருகின்றதென்றால், அந்த மதிப்பு 1.96யை விட அதிகமாக உள்ளது. அதாவது, இரண்டு கூட்டுச்

சராசரிகளுக்கிடையே உள்ள வித்தியாசம் அவற்றின் திட்டப் பிழையைப்போல் 3.2 மடங்கு அதிகம் என்று பொருள். வித்தியாசம் 1.96 மடங்குக்கும் மேல் இருந்தால் வித்தியாசம் அதிகம் என்று சொல்லும்போது, நாம் சொல்வது சரியாக இருக்க 95 சதவீத வாய்ப்பு உள்ளது; தவறாக இருப்பதற்கான வாய்ப்பு 5 சதவீதமே. z மதிப்பு மிக முக்கிய பங்கினை வகிப்பதால் இதற்கு 'சோதனைப்புள்ளி' (test statistic) என்ற பெயரும் உண்டு.

எடுத்துக்கொண்டுள்ள கருத்தை விளக்குவதற்கு இன்னுமொரு உதாரணம் எடுக்கலாம். உதாரணமாக, ஒரு முழுமையின் (population) சராசரி 100 ஆகவும், அதன் திட்டவிலக்கம் 24 ஆகவும் உள்ளதெனக் கொள்வோம். அந்த முழுமையிலிருந்து ஒரு மாதிரி 64 உறுப்புக்களுடன் எடுக்கப்படுகிறது. அந்த மாதிரியின் சராசரி 110 ஆக உள்ளது. இப்படிப்பட்ட புள்ளி விபரங்களை வைத்து அந்த முழுமையின் சராசரிக்கும் மாதிரியின் சராசரிக்கும் உள்ள வித்தியாசம் புள்ளியியல் முக்கியத்துவம் பெற்றதா, இல்லையா என்று பார்க்கலாம்.

முதலில் எடுகோள்களை உருவாக்கிக் கொள்ள வேண்டும்.

$$H_0 : \bar{X} = \mu$$

$$H_1 : \bar{X} \neq \mu$$

முழுமைக்கும் மாதிரிக்கும் உள்ள தொடர்புகளில் ஒன்று, முழுமையின் சராசரிக்குப் பக்கத்தில்தான் மாதிரிகளின் சராசரியும் இருக்கும் என்பது ஆகும். அப்படியின்றி, வெகுதூரத்தில் தள்ளியிருந்தால் எந்தெந்த மாதிரிகளின் சராசரி தள்ளி உள்ளனவோ, அந்த மாதிரிகள் கொடுக்கப்பட்டுள்ள முழுமையிலிருந்து எடுக்கப்படாமல் வேறு ஒரு முழுமையிலிருந்து எடுக்கப்பட்டிருக்கலாம்.

அல்லது, அந்த மாதிரிகள் சமவாய்ப்பு முறைப்படி எடுக்கப்படாமல் இருக்கலாம். சமவாய்ப்பு முறைப்படி கொடுக்கப்பட்டுள்ள முழுமையிலிருந்து மாதிரி எடுக்கப் பட்டிருந்தால், அந்த மாதிரியின் சராசரி ($\overline{X}$) முழுமையின் சராசரிக்கு (μ) அருகில்தான் இருக்கும்; ஏனெனில் மிகச்சில மாதிரிகளின் சராசரிகள்தான் தூரத்தில் இருக்கும். இதுவரை தெரிந்துள்ளபடி, 95 சதவீத மாதிரிகளின் கூட்டுச்சராசரிக்கும் முழுமையின் கூட்டுச்சராசரிக்கும் உள்ள வித்தியாசம் திட்டப் பிழையைப் போல் 1.96 மடங்குகள் இருக்கும்; 99 சதவீத மாதிரிகளின் கூட்டுச் சராசரிக்கும் முழுமையின் கூட்டுச் சராசரிக்கும் உள்ள வித்தியாசம் திட்டப்பிழையைப் போல் 2.58 மடங்குக்குள் இருக்கும்.

கிடைத்துள்ள விபரங்களின்படி zயைக் கணக்கிட்டால்,

$$z_c = \frac{\overline{X} - \mu}{SE} = \frac{110 - 100}{24/\sqrt{64}} = \frac{10}{3} = 3.33$$

கணக்கிடப்பட்ட z, 2.58யைவிட அதிகமாக உள்ளது. எனவே அது வரைபடம் 38இல் கொடுக்கப்பட்டுள்ள நிழலாக்கிய பகுதியில் வருகிறது. அப்படியானால், இல்லெனும் எடுகோளை நிராகரிக்கின்றோம். அதாவது $\overline{X}$க்கும் μக்கும் உள்ள வித்தியாசம் புள்ளியியல் முக்கியத்துவம் வாய்ந்தது என்கிறோம். இப்படிச் சொல்வது சரியாக இருப்பதற்கான நிகழ்தகவு 0.99; தவறாக இருப்பதற்கான நிகழ்தகவு 0.01 தான்.

z_cன் மதிப்பு பகுதியில் உள்ள வித்தியாசத்தின் அளவையும் கீழே தொகுதியில் உள்ள திட்டப்பிழையையும் பொறுத்துள்ளது. திட்டப்பிழை திட்டவிலக்கத்தையும் மாதிரியின் அளவையும் பொறுத்துள்ளது. z_cக்கும் பகுதியில் உள்ள மதிப்புக்கும் நேரிடை உறவு உள்ளது. z_cக்கும் திட்டப் பிழைக்கும் எதிரிடை உறவு உள்ளது.

மேலே கூறப்பட்டுள்ள உதாரணத்தில் $\overline{X}$ன் மதிப்பு 105ஆகக் (குறைவாக) இருந்திருந்தால், மாதிரியின் கூட்டுச் சராசரிக்கும் முழுமையின் கூட்டுச் சராசரிக்கும் உள்ள வித்தியாசம் குறைவாக இருக்கும். இப்போது z_cன் மதிப்பு $\frac{5}{3} = 1.66$ தான். இது மிகக் குறைவான வித்தியாசம். எனவே, இல்லெனும் எடுகோள் ஒத்துக் கொள்ளப்படுகிறது. இந்த முடிவு சரியாக இருக்க வாய்ப்பு 95 சதவீதம், தவறாகப் போவதற்குள்ள வாய்ப்பு 5 சதவீதம்தான். இந்த முடிவு, 5 சதவீத புள்ளியியல் முக்கியத்துவம் பெற்றுள்ளதாகக் கருதப்படுகிறது.

ஒருபுறச் சோதனையும் இருபுறச் சோதனையும் (ONE - TAILED TEST and TWO-TAILED TEST)

இருபுறச் சோதனையில் வித்தியாசம் இருக்கிறதா, இல்லையா என்பதுதான் தெரியும். கூடவா, குறையவா என்பது தெரியாது. சில சமயம் இது முக்கியம். மணமகனின் உயரத்திற்கும் மணமகளின் உயரத்திற்கும் வித்தியாசம் உள்ளது என்று சொன்னால், பொதுவாக மணமகன் மணமகளைவிட உயரம் என்றுதான் பலரும் நினைப்பார்கள். ஆனால், கொடுக்கப்பட்டுள்ள வாக்கியத்திற்கு அது மட்டும் பொருள் அல்ல. மணமகள் மணமகனை விட உயரமாக இருந்தாலும் மணமகனின் உயரத்திற்கும் மணமகளின் உயரத்திற்கும் வித்தியாசம் உண்டு என்று சொல்லலாம். இப்படிப்பட்ட குழப்பங்களைத் தவிர்க்க, $\mu_1 \neq \mu_2$ என்று சொல்வதைவிட, μ_1, μ_2வை விடப் பெரியதா, μ_2, μ_1யைவிடப் பெரியதா என்று சொல்வது நல்லது. $\mu_1 < \mu_2$ அல்லது $\mu_2 < \mu_1$ அல்லது $\mu_1 > \mu_2$ என்று சொல்வது ஒருபுறச் சோதனையாகும். $\mu_1 \neq \mu_2$ என்பது, இருபுறச் சோதனையாகும்.

ஒருபுறச் சோதனையில், நிராகரிக்கின்ற பரப்பளவை (critical region) ஒரு புறம் மட்டும் வைக்கின்றோம். உதாரணமாக, $\mu < 100$ என்று சொல்லும்போது நிராகரிக்கின்ற

பரப்பளவு முழுவதும் இடப்புறமாகவே இருக்கும். மாறாக $\mu > 100$ என்று சொல்லும்போது நிராகரிக்கின்ற பரப்பளவு முழுதும் வலதுபுறமாகவே இருக்கும். நிராகரிக்கின்ற பகுதி 5 சதவீதம் என்றால், 5 சதவீதப் பரப்பும் ஒருபுறமாகவே இருக்கும். இவற்றிற்கான செய்திகள் அட்டவணை 48இல் தரப்பட்டுள்ளன.

அட்டவணை - 48

முக்கியத்துவத்தின் நிலை $\propto$	0.10	0.05	0.01	0.005	0.002
மறுத்தலுக்கான z ன் மதிப்புக்கள் (ஒருபுறச் சோதனை)	-1.28 (or) 1.28	-1.645 (or) 1.645	-2.33 (or) 2.33	-2.58 (or) 2.58	-2.88 (or) 2.88
மறுத்தலுக்கான z ன் மதிப்புக்கள் (இருபுறச்சோதனை)	-1.645 and 1.645	-1.96 and 1.96	-2.58 and 2.58	-2.81 and 2.81	-3.08 and 3.08

சிறப்புச் சோதனைகள் (Special tests)

பெரிய மாதிரிகளுக்கு (large sample), மாதிரிகளின் பரவல்கள் (sampling distributions) இயல்நிலைப் பரவல்களாக (normal distributions) இருக்கும்.

அட்டவணை - 49

முடிவில்லா முழுமை அல்லது திரும்பச் சேர்த்து மாதிரி எடுத்தல்	திரும்பச் சேர்க்காமல், முடிவுள்ள முழுமையிலிருந்து மாதிரி எடுத்தல்
1. $z = \dfrac{\overline{x}-\mu}{\sigma/\sqrt{n}}$	$\sigma/\sqrt{n}$க்குப் பதிலாக $\left(\dfrac{\sigma}{\sqrt{n}}\right)\sqrt{\dfrac{N-n}{N-1}}$
2. $z = \dfrac{p-P}{\sqrt{PQ/n}}$	$\sqrt{PQ/n}$க்குப் பதிலாக $\left(\dfrac{\sqrt{PQ}}{\sqrt{n}}\right)\sqrt{\dfrac{N-n}{N-1}}$

3. $z = \frac{X - nP}{\sqrt{nPQ}}$ இதில் $p = \frac{x}{n}$, $\mu = nP$ $s = \sqrt{nPQ}$	—
4. $z = \frac{\bar{X}_1 - \bar{X}_2 - 0}{\sqrt{\left(\frac{\sigma_1^2}{n_1}\right) + \left(\frac{\sigma_2^2}{n_2}\right)}}$ σ_1, σ_2க்களின் மதிப்புகளாக s_1, s_2அல்லது $\hat{s}_1$, $\hat{s}_2$க்கள் பயன்படுத்தப்படலாம்	Standard error becomes $\sqrt{\left(\frac{\sigma_1^2}{n_1}\right)\left(\frac{N_1 - n_1}{N_1 - 1}\right) + \left(\frac{\sigma_2^2}{n_2}\right)\left(\frac{N_2 - n_2}{N_2 - 1}\right)}$
5. $z = \frac{p_1 - p_2 - 0}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$ இதில் $P = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$	Standard error becomes $\sqrt{\frac{P_1 Q_1}{n_1} + \frac{P_2 Q_2}{n_2}}$

சிறிய மாதிரிகளின் கோட்பாடு

(Small Sampling Theory) ('t' மற்றும் χ^2 பரவல்கள்)

முழுமையும் மாதிரிகளும் இயல்நிலைப் பரவலைக் கொண்டிருக்கும் என்ற அநுமானத்துடன் இதற்கு முந்தைய சோதனைகள் விவரிக்கப்பட்டன. முழுமை முடிவில்லாமல் அல்லது மிகப்பெரியதாக இருந்தாலோ, மாதிரிகள் மிகப் பெரியதாக இருந்தாலோ, மாதிரிகளின் கூறுகள் திரும்பத் திரும்ப முழுமையுடன் சேர்க்கப்பட்டு மாதிரிகள் சமவாய்ப்பு முறைப்படி எடுக்கப்பட்டாலோ மாதிரியின் பரவல்கள் இயல்நிலைப் பரவலாக அல்லது இயல்நிலைப் பரவலை ஒத்து இருக்கும். ஆனால், மேலே கூறப்பட்ட சூழ்நிலைகள்

இல்லாமலும் சில சமயங்களில் இருக்கலாம். அப்படி இருக்கும்போது வேறு சில சோதனைகளை மேற்கொள்ளலாம்.

மாதிரியின் அளவு பெரியதாக (n > 30) இருந்தால் இயல்நிலைப் பரவலுக்கான சோதனையைப் பயன்படுத்தலாம். சிறிய மாதிரியாக (small sample) இருப்பதற்கான சூழ்நிலைகளும் உள்ளன. மாதிரியின் ஒவ்வொரு கூறினையும் தேர்ந்தெடுக்கப் பணம் நிறையச் செலவாகலாம்; அல்லது, அதிகக் காலம் பிடிக்கலாம். மாதிரிகளுக்கான கூறுகளை எடுக்கும்போதே அதன் உபயோகம் குறைந்து விடலாம். இதுபோன்ற சூழ்நிலைகளில் சிறிய மாதிரியையே எடுத்து அந்த மாதிரி எடுக்கப்பட்ட முழுமையைப் பற்றி அறிய வேண்டியுள்ளது. இவ்வித சிறிய மாதிரிகளுக்குத் திருத்தப்பட்ட மாதிரிகள் (exact samples) என்றொரு பெயரும் உண்டு.

மாணவரின் ‘t’ பரவல்கள் (Student’s ‘t’ distribution)

முழுமையின் திட்டவிலக்கம் தெரியாதபோதும், மாதிரியின் அளவு சிறியதாக (n < 50 என்று டேரோ யமனேயின் முன்னர் கூறப்பட்டுள்ள புத்தகத்தில் 647ம் பக்கத்தில் உள்ளது. சில புத்தகங்களில் n < 30 என்றும் உள்ளது) உள்ளபோதும் நிலைப்படுத்தப்பட்ட z இயல்நிலைப் பரவலைக் கொண்டிருக்காது. இந்தச் சூழலுக்குப் பொருத்தமான பரவலை கோஸ்ஸெட் (W.S.Gosset, ‘The Probable Error of a Mean’, Biometrika, 1908) என்பவர் ‘மாணவர்’ என்ற புனைப்பெயரில், 1908இல் தந்துள்ளார். அதுவே ‘t’ பரவல் என அழைக்கப்படுகிறது. அதன்படி

$$t = \frac{\bar{x} - \mu}{\hat{s}/\sqrt{n}}$$

மாதிரிகளின் அளவு கூடி 30க்கும் மேல் வரும்போது ‘t’ பரவல் இயல்நிலைப் பரவலுடைய குணங்களைப் பெறும். இயல்நிலைப் பரவல் மாதிரிகளின் அளவால் (30க்கும் மேல்) பாதிக்கப்படாது. ஆனால், ‘t’ சிறிய மாதிரியைப் பற்றியதால் t’ன் பரவல் மாதிரியின் அளவால் பாதிக்கப்படுகிறது. மாதிரியின் அளவு குறையக் குறையத் தட்டைத்தன்மை கூடிக் கொண்டும், மாதிரியின் அளவு கூடக்கூட தட்டைத் தன்மை குறைந்து இயல்நிலைப் பரவல் போன்றும் மாறும்.

இயல்நிலைப் பரவலுக்குச் செய்தது போலவே, ‘t’ பரவலைப் பயன்படுத்தி, 95%, 99% நம்பிக்கை இடைவெளிகளைக் காணலாம். உதாரணமாக, முழுமையின் கூட்டுச்சராசரி $\bar{X} + t_{0.975} \frac{s}{\sqrt{n-1}}$ யை விடக் குறைவாகவும் $\bar{X} - t_{0.975} \frac{s}{\sqrt{n-1}}$ யை விடக் கூடுதலாகவும் இருக்கும் என்று 95 சதவீத நம்பிக்கையுடன் சொல்லலாம்.

பொதுவாக, முழுமையின் கூட்டுச் சராசரியின் நம்பிக்கை எல்லைகளை $\bar{X} \pm t_c \frac{s}{\sqrt{n-1}}$ என்று சொல்லலாம். இதில் t_c என்பது நம்பிக்கைக் கெழு (confidence coefficient). நம்பிக்கைக்கெழு தேவையான நம்பிக்கையின் அளவையும் (level of confidence), மாதிரியின் அளவையும் (sample size) பொறுத்து அமைகிறது.

$$t = \frac{\bar{x}-\mu}{s}\sqrt{n-1} = \frac{\bar{x}-\mu}{\hat{s}}\sqrt{n}$$

$$\hat{s} = s\sqrt{\frac{n}{n-1}}$$

உதாரணமாக, n_1 என்ற சமவாய்ப்பு மாதிரி, இயல்நிலைப் பரவலைக் கொண்ட ஒரு முழுமையிலும், n_2 என்ற சமவாய்ப்பு மாதிரி இயல்நிலைப் பரவலைக்கொண்ட இன்னொரு

முழுமையிலிருந்தும் சார்பிலா (independent) முறையில் எடுக்கப்படுவதாகவும், இரண்டு முழுமைகளின் திட்ட விலக்கங்களும் ($\sigma_1 = \sigma_2$) சமமாக இருப்பதாகவும் கொள்வோம். மேலும் இந்த இரு மாதிரிகளும் $\overline{X}_1$, $\overline{X}_2$ கூட்டுச் சராசரிகளையும் s_1, s_2 என்ற திட்டவிலக்கங்களையும் கொண்டுள்ளதாகவும் கொள்வோம். இப்படிப்பட்ட சூழலில் அந்த இரண்டு மாதிரிகளும் ஒரே முழுமையிலிருந்துதான் எடுக்கப்பட்டு உள்ளனவா என்று சோதனை செய்ய:

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$ எனும் சூத்திரத்தை எடுத்துக் கொள்ளலாம்.

இதில்,

$$\sigma = \sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}}$$

$n_1 + n_2 - 2$ என்பது கட்டின்மை எண்ணிக்கை (degrees of freedom) ஆகும்.

சார்பில்லா மாதிரிகளாக இருந்தால்,

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

எனும் சூத்திரத்தைப் பயன்படுத்தி $\mu_1 = \mu_2$ என்ற இல்லெனும் எடுகோளைச் சோதனை செய்யலாம். மாறாகச் சார்புள்ள மாதிரிகளாக இருந்தால், வேறு ஒரு வழியினைக் கடைப்பிடித்து μ_1க்கும் μ_2க்கும் வித்தியாசம் இல்லை எனும் இல்லெனும் எடுகோளைச் சோதனை செய்ய வேண்டும். உதாரணமாக, ஒரு போட்டித் தேர்வில் 10 மாணவர்கள் பெற்ற மதிப்பெண்கள் (நூற்றுக்கு) கொடுக்கப்பட்டுள்ளன. அதே 10 மாணவர்கள் ஒரு கல்லூரியில் சேர்ந்து 10 மாதங்கள் படித்து

விட்டு முன்னால் எழுதிய அளவிலான மற்றொரு போட்டித் தேர்வை எழுதி, அதில் பெற்ற மதிப்பெண்களும் (நூற்றுக்கு) கொடுக்கப்பட்டுள்ளன. இந்த இரண்டு தேர்வுகளிலும் மாணவர்கள் பெற்றுள்ள மதிப்பெண்களை ஒப்பிட்டுப் பார்த்துவிட்டு கல்லூரியில் சேர்ந்து படித்துவிட்டுப் பின்னர் எழுதிய தேர்வில் பெற்ற மதிப்பெண்கள் குறைந்துள்ளதா என்றும், அப்படிக் குறைந்து இருந்தால் அது புள்ளியியல் முக்கியத்துவம் பெற்றதா என்றும் காணவிரும்பலாம். இச்சூழலில் இது சார்பிலா மாதிரியல்ல. அதே மாணவர்கள் இரண்டு தேர்வுகளிலும் பெற்ற மதிப்பெண்கள் சார்ந்திருக்கலாம். அப்படிப்பட்ட சூழலில் : $t = \frac{\bar{d}\sqrt{n}}{s}$ இதில் $\bar{d} = \frac{\Sigma d}{n}$ $s = \frac{\Sigma d^2 - n(\bar{d})^2}{n-1}$; d என்பது இரண்டு தேர்வுகளிலும் ஒவ்வொரு மாணவரும் பெற்ற மதிப்பெண்களுக்கிடையேயான வித்தியாசம் (difference). அட்டவணை 49 (அ) இந்த விபரங்களைத் தருகிறது.

அட்டவணை - 49(அ)

முதலில் எழுதிய தேர்வில் பெற்ற மதிப்பெண்கள்	80	76	92	60	70	56	74	56	70	56
கல்லூரியில் படித்த பிறகு எழுதிய தேர்வில் பெற்ற மதிப்பெண்கள்	70	72	80	50	60	48	70	46	60	40
Difference 1st - 2nd	10	4	12	10	10	8	4	10	10	16
d^2	100	16	144	100	100	64	16	100	100	256

H_0 : இரண்டு தேர்வுகளிலும் பெற்ற மதிப்பெண்களுக்கு இடையே வித்தியாசம் இல்லை.

(அல்லது)

முதல் தேர்வில் பெற்ற மதிப்பெண்களின் சராசரியும் (μ_1) இரண்டாவது தேர்வில் பெற்ற மதிப்பெண்களின் சராசரியும் (μ_2) சமம் ($\mu_1 = \mu_2$)

H_1 : முதலில் பெற்ற மதிப்பெண்களின் சராசரி (μ_1) இரண்டாவதாகப் பெற்ற மதிப்பெண்களின் சராசரியை விட (μ_2) அதிகம் ($\mu_1 > \mu_2$)

$$\boxed{t = \frac{\overline{d}\sqrt{n}}{s}} \qquad \overline{d} = \frac{\Sigma d}{n} = \frac{94}{10} = \boxed{9.4}$$

$\Sigma d = 94$

$$s = \sqrt{\frac{\Sigma d^2 - n(\overline{d})^2}{n-1}}$$

$\overline{d}^2 = 88.36$

$$= \sqrt{\frac{906 - 883.6}{9}} = \boxed{2.49}$$

$\Sigma d^2 = 906$

$n = 10$

$$t = \frac{9.4\sqrt{10}}{2.49}$$

$$= \frac{29.73}{2.49} = \boxed{11.94}$$

't' பட்டியலிலிருந்து கிடைக்கும் மதிப்பு ($\gamma = 10\text{-}1 = 9$; $t_{0.05}$) = 2.62. கணக்கிடப்பட்ட 't' ன் மதிப்பு பட்டியலிடப்பட்ட tன் மதிப்பைவிட அதிகம். எனவே இல்லெனும் எடுகோள் ஒத்துக்கொள்ளப்படவில்லை; மாற்று எடுகோள் ஆமோதிக்கப்படுகிறது. எனவே, கல்லூரியில் சேர்ந்து படித்த பின்னர் எழுதிய போட்டித் தேர்வில் மதிப்பெண்கள் குறைந்துள்ளன எனலாம்.

கட்டின்மை எண்ணிக்கை (degrees of freedom)

கட்டின்மை எண்ணிக்கை என்பது கிரேக்க (Greek) எழுத்தான v (nu) என்பதால் குறிக்கப்படுகிறது. இந்த v = n-k. இதில் n என்பது மாதிரியில் உள்ள சார்பிலாக் கூறுகள். K என்பது மாதிரியை வைத்து மதிப்பீடு செய்யப்பட வேண்டிய முழுமையின் பண்பலகுகள்.

இன்னும் எளிமையாகச் சொல்ல வேண்டுமானால் 5 எண்கள் இருந்தால் கட்டின்மை எண்ணிக்கை 4 ஆகும். உதாரணாக 5 எண்களின் கூடுதல் 25 என்றால் 4 எண்களை சேர்த்துக் கொள்ள எந்த கட்டுப்பாடும் இல்லை; எந்த நான்கு எண்களையும் சேர்த்துக் கொள்ளலாம். ஆனால், ஐந்தாவது எண்ணைச் சேர்ப்பதற்கு கட்டுப்பாடு இருக்கிறது. எடுத்துள்ள உதாரணத்தில் முதல் நான்கு எண்கள் 10, 20, 5, 3 என எதை வேண்டுமானாலும் சேர்த்துக் கொள்ளலாம். இந்த நான்கு எண்களின் கூடுதல் 38 வருகிறது. எனவே, ஐந்தாவது எண் - 13 ஆகத்தான் இருக்கவேண்டுமென்ற கட்டுப்பாட்டினை மொத்தமான 25 விதிக்கிறது. எனவே இதில் n-1 தான் (5-1=4) கட்டின்மையின் எண்ணிக்கை.

ஓர் அட்டவணைக்கு எவ்வாறு கட்டின்மை எண்ணிக்கை கண்டுபிடிப்பது? இரண்டு நிரல்களும் (Columns) இரண்டு நிரைகளும் (Rows) இருந்தால் கட்டின்மை எண்ணிக்கை (நிரல் - 1) (நிரை - 1) = (2-1) (2-1) = 1. நிரல்கள் 3 ஆகவும் நிரைகள் 3 ஆகவும் இருந்தால் கட்டின்மை எண்ணிக்கை (3-1) (3-1) = 4 ஆகும். உதாரணமாக அட்டவணை 50ஐப் பார்க்கலாம். இது ஒரு 2 x 2 அட்டவணை. இதற்கு கட்டின்மை எண்ணிக்கை (2-1) (2-1) = 1 தான். இது எப்படி என்று அட்டவணை 50லிருந்து புரிந்து கொள்ளலாம். இந்த அட்டவணையில் உள்ள நான்கு அறைகளில் (Cells) ஓர் அறையை ஓர் எண்ணால் பூர்த்தி செய்வதற்குத்தான் சுதந்திரம்

உண்டு. ஏதேனும் ஓர் எண்ணை எழுதிவிட்டால் மற்ற மூன்று எண்களையும் விருப்பப்படி எழுத முடியாது. எனவே, கட்டின்மை எண்ணிக்கை இதற்கு ஒன்றுதான்.

அட்டவணை - 50

	படித்தவர்கள்	படிக்காதவர்கள்	மொத்தம்
பண்புள்ளவர்கள்			150
பண்பில்லாதவர்கள்			50
மொத்தம்	100	100	200

χ^2 சோதனை (χ^2 Test)

χ^2 என்பது கிரேக்க எழுத்தான Chiயிலிருந்து வந்தது. இந்தச் சோதனை பலவிதமாகப் பயன்படுகிறது. பெரிய மாதிரிக்கும் சிறிய மாதிரிக்கும் இதைப் பயன்படுத்தலாம். இதை ஒரு பண்பலகுச் சோதனையாகவும் (Parametric test) பண்பலகில்லாச் சோதனையாகவும் (non-parameteric test) பயன்படுத்தலாம். இது மேலும், கோட்பாட்டுப் பரவலின் ஒத்த தன்மையை (goodness of fit) சோதிக்கவும், சார்பிலாத்தன்மையைச் (independence) சோதிக்கவும், சமபடித்தன்மையை (homgeneity) சோதிக்கவும் பயன்படுகிறது.

இந்த χ^2 சோதனையைப் பயன்படுத்தி முழுமையின் மாறுபாட்டுக்கும் (variance) மாதிரியின் மாறுபாட்டுக்கும் உள்ள வித்தியாசம் புள்ளியியல் முக்கியத்துவம் பெற்றதா என்று கண்டுபிடிக்கலாம். அதற்கான சூத்திரம்:

$$\chi^2 = \frac{ns^2}{\sigma^2}$$

இந்த மதிப்பு χ^2 பட்டியலில் கிடைக்கும் மதிப்பைவிட அதிகமாக இருந்தால், s^2க்கும் σ^2க்கும் உள்ள வித்தியாசம் புள்ளியியல் முக்கியத்துவம் வாய்ந்தது என்று பொருள்.

அல்லது s^2 கிடைத்த மாதிரி, σ^2 கிடைத்த முழுமையிலிருந்து எடுக்கப்படவில்லை என முடிவு செய்யலாம்.

இந்தப் பரவலைப் பயன்படுத்தி முழுமையின் திட்டவிலக்கம் எவ்வளவுக்குள் இருக்கும் என்று நம்பிக்கை இடைவெளி காணலாம். உதாரணமாக,

$$\frac{s\sqrt{n}}{\chi_{0.975}} < \sigma < \frac{s\sqrt{n}}{\chi_{0.025}}$$

என்று 95 சதவீத நம்பிக்கையுடன் சொல்லலாம்.

கிடைத்த மற்றும் கோட்பாட்டு அலைவெண்கள் (Observed and theoretical frequencies)

கோட்பாட்டுப்படி எதிர்பார்க்கின்ற விளைவுகளுக்கும் உண்மையாக நடக்கின்ற விளைவுகளுக்கும் தொடர்பு இருக்கும் என்று சொல்லலாம் என்றாலும், ஒவ்வொரு விளைவுக்கும் அந்தத் தொடர்பு வெளிப்படையாகத் தெரியாமலும் இருக்கலாம். 100 தடவைகள் ஒரு பேதமற்ற நாணயத்தைச் சுண்டினால் 50 தடவைகள் பூவும் 50 தடவைகள் தலையும் மேலே தெரிய வேண்டும் என்று ஈருறுப்புக் கோட்பாடு சொல்கிறது. ஆனாலும், சரியாக 50 தடவைகள் பூவும் 50 தடவைகள் தலையும் வருமா என்றால் சந்தேகம்தான். ஓர் ஆறு சமபக்கங்கள் கொண்ட பகடைக் கட்டையை அறுபது தடவைகள் உருட்டினால் 1, 2, 3, 4, 5, 6 ஆகிய ஒவ்வொன்றும் 10 தடவைகள் வர வேண்டும் என்கிறது கோட்பாடு. ஆனால், உண்மையில் அவ்வாறு நிகழுமா என்பது நிச்சயமல்ல. இப்படிப்பட்ட சூழ்நிலைகளில்,

$$\chi^2_c = \Sigma\ \frac{(o-e)^2}{e}$$

என்ற சூத்திரத்தைப் பயன்படுத்தி, கணக்கிடப்பட்ட χ^2 பட்டியலில் உள்ள χ^2 மதிப்பைவிட அதிகமாக இருந்தால், கிடைக்கப்பெற்ற அலைவெண்களுக்கும் கோட்பாட்டுப்படி

எதிர்பார்க்கப்படும் அலைவெண்களுக்கும் உள்ள வித்தியாசம் புள்ளியியல் முக்கியத்துவம் பெற்றதாகும். அப்படியானால், கோட்பாடும் உண்மை நிலவரமும் ஒத்துப் போகவில்லை என்று பொருள். அப்படியானால், முயற்சிகளின் (trials) எண்ணிக்கையைக் கூட்டவேண்டும். அப்படியும் மேலே கூறப்பட்ட வித்தியாசம் அதிகமாக இருந்தால், அந்தப் பகடைக்கட்டையில் ஏதோ பிரச்சனை உண்டு என்று பொருள்.

யேட்ஸ் உடைய திருத்தம் (YATES' CORRECTION)

தொடர் பரவல்களின் (continuous distributions) முடிவுகளை தனித்த விபரங்களுக்கு (discrete data) பயன்படுத்தும்போது, தொடர்ச்சிக்காகச் சில திருத்தங்கள் செய்ய வேண்டியுள்ளது. χ^2 பரவலுக்கும் அதுபோன்ற திருத்தங்கள் இருக்கின்றன.

திருத்தப்பட்ட χ^2

$$= \frac{(|o_1-e_1|-0.5)^2}{e_1} + \frac{(|o_2-e_2|-0.5)^2}{e_2} + \ldots + \frac{(|o_K-e_K|-0.5)^2}{e_K}$$

இது யேட்ஸ் உடைய திருத்தம் என அழைக்கப்படுகிறது. யேட்ஸ் திருத்தம் எப்போதும் χ^2மதிப்பைக் குறையச் செய்யும். இதுபோன்ற திருத்தங்கள் கட்டின்மை எண்ணிக்கை 1 ஆக இருக்கும்போதோ, மாதிரி மிகச் சிறியதாக இருக்கும்போதோ, எதிர்பார்க்கும் அலைவெண் மிகச்சிறியதாக (10க்கும் குறைவாக) இருக்கும்போதோ தேவைப்படுகின்றது. மேலேகூறப்பட்டுள்ள காரணங்களில் ஏதோ ஒன்றாலோ, பலவற்றாலோ, திருத்தப்பட்ட χ^2யினால் கிடைக்கும் முடிவுக்கும் திருத்தப்படாத χ^2யினால் கிடைக்கும் முடிவுக்கும் இடையே வித்தியாசம் இருந்தால் மாதிரியின் அளவைக் கூட்ட வேண்டும்.

χ^2 கணிப்பதற்கு எளிய சூத்திரம்

கோட்பாட்டுப்படி எதிர்பார்க்கப்படும் அலைவெண்கள் இல்லாத சூழ்நிலையில், சேகரிக்கப்பட்ட அலைவெண்களை மட்டும் வைத்து χ^2யைக் கணிப்பதற்கான சூத்திரம் முதலில் 2 x 2 அட்டவணைக்கு (அட்டவணை 51இல்) தரப்படுகிறது.

$$\chi^2 = \frac{N\Delta^2}{N_1N_2N_AN_B}$$

$\Delta = a_1b_2 - a_2b_1$

$N = a_1 + a_2 + b_1 + b_2$ அல்லது $N_1 + N_2$ அல்லது $N_A + N_B$

$N_1 = a_1 + b_1$; $N_2 = a_2 + b_2$

$N_A = a_1 + a_2$; $N_B = b_1 + b_2$

அட்டவணை - 51

	I	II	மொத்தம்
A	a_1	a_2	N_A
B	b_1	b_2	N_B
மொத்தம்	N_1	N_2	N

யேட்ஸ் திருத்தத்துடன் 2 x 2 அட்டவணைக்கான (அட்டவணை 51) χ^2 :

$$\chi^2 = \frac{N(|\Delta| - \frac{1}{2}N)^2}{N_1N_2N_AN_B}$$

2 x 3 அட்டவணைக்கான χ^2 (அட்டவணை 52யைப் பார்க்கவும்)

$$\chi^2 = \frac{N}{N_A}\left[\frac{a_1^2}{N_1} + \frac{a_2^2}{N_2} + \frac{a_3^2}{N_3}\right] + \frac{N}{N_B}\left[\frac{b_1^2}{N_1} + \frac{b_2^2}{N_2} + \frac{b_3^2}{N_3}\right] - N$$

அட்டவணை - 52

	I	II	III	மொத்தம்
A	a_1	a_2	a_3	N_A
B	b_1	b_2	b_3	N_B
மொத்தம்	N_1	N_2	N_3	N

இப்போது, கீழே உள்ள கணக்கினைச் செய்து பார்க்கலாம். மூன்று தேர்வாளர்கள், அன்பு, ஆதி, இசை ஆகியோர் நடத்திய தேர்வில் தேர்ச்சி பெற்ற, தேர்ச்சி பெறாத மாணவர்களின் எண்ணிக்கைகள் அட்டவணை 53இல் கொடுக்கப்பட்டுள்ளளன. மூன்று தேர்வாளர்களிடமும் தேர்ச்சி பெறாதவர்களின் விகிதாச்சாரம் (proportion) சமமாக உள்ளதா எனக் காணலாம்.

அட்டவணை - 53
சேகரிக்கப்பட்ட அலைவெண்கள்

	அன்பு	ஆதி	இசை	மொத்தம்
தேறியவர்கள்	50	47	56	153
தேறாதவர்கள்	5	14	8	27
மொத்தம்	**55**	**61**	**64**	**180**

முதலில் இல்லெனும் எடுகோளைக் கொடுக்கலாம்.

H_0 : மூன்று தேர்வாளர்களிடமும் தேர்ச்சி பெறாதவர்களின் விகிதாச்சாரம் சமம்.

இதில் மொத்தத்தில் தேர்ச்சி பெறாதவர்களின் விகிதம் : $27 \div 180 = 15$ சதவீதம். இதேபோல், அன்பு, ஆதி, இசை மூவரும் தேர்வு முடிவுகள் கொடுத்திருந்தால், 8.25 மாணவர்கள் அன்பிடமும், 9.15 மாணவர்கள் ஆதியிடமும் 9.60 மாணவர்கள் இசையிடமும் தேர்ச்சி பெற்றிருக்க

மாட்டார்கள். இந்த விபரங்களைக் கொண்டு, எதிர்பார்க்கப்படும் நிகழ்வெண்களைக் கண்டுபிடிக்கலாம். அவை அட்டவணை 54இல் தரப்பட்டுள்ளன. தேர்ச்சி அடையாதவர்களின் எண்ணிக்கையிலிருந்து தேர்ச்சி பெற்றவர்களின் எண்ணிக்கையையும் (85 சதவீதம்) கணக்கிடலாம். அவை முறையே 46.75, 51.85, 54.40 ஆகும்.

அட்டவணை - 54
எதிர்பார்க்கப்படும் அலைவெண்கள்

	அன்பு	ஆதி	இசை	மொத்தம்
தேர்ந்தவர்களின் எண்ணிக்கை	55ன் 85 சதவீதம் 46.75	61ன் 85 சதவீதம் 51.85	64ன் 85 சதவீதம் 54.40	153
தேராதவர்களின் எண்ணிக்கை	55ன் 15 சதவீதம் 8.25	61ன் 15 சதவீதம் 9.15	64ன் 15 சதவீதம் 9.60	27
மொத்தம்	55	61	64	180

அட்டவணை 54ன் உள்கட்டங்களில் உள்ள எண்களை வேறுவிதமாகவும் கண்டுபிடிக்கலாம்.

$$e_{11} = \frac{153 \times 55}{180} = 46.75 = \frac{(\text{நிரை மொத்தம்} \times \text{நிரல் மொத்தம்})}{(\text{முழு மொத்தம்})}$$

$$e_{12} = \frac{153 \times 61}{180} = 51.85 = \quad (\text{அதே})$$

$$e_{13} = \frac{153 \times 55}{180} = 54.40 = \quad (\text{அதே})$$

$$e_{21} = \frac{27 \times 55}{180} = 8.25 = \quad (\text{அதே})$$

$$e_{22} = \frac{27 \times 61}{180} = 9.15 = \quad (அதே)$$

$$e_{23} = \frac{27 \times 64}{180} = 9.60 = \quad (அதே)$$

$$\chi^2 = \frac{(50 - 46.75)^2}{46.75} + \ldots + \frac{(8 - 9.60)^2}{9.60} = 4.84$$

அட்டவணை 54க்கு கட்டின்மை எண்ணிக்கை (2-1) (3-1) = 2 ஆகும். அதாவது, உள்கட்டத்தில் உள்ள 2 நிரை 3 நிரல்களில் உள்ள ஆறு கட்டங்களில் ஏதேனும் 2 கட்டங்களில் எண்கள் எழுத கட்டின்மை (freedom) உள்ளது. ஏதேனும் இரண்டு எண்களைப் பூர்த்தி செய்துவிட்டால் மற்ற எண்களைப் பூர்த்தி செய்யக் கட்டுப்பாடு உள்ளது. எனவே, இந்த அட்டவணைக்குக் கட்டின்மை 2.

χ^2பட்டியலின்படி, 95 சதவீத நம்பிக்கை அளவில், 5 சதவீத முக்கியத்துவ அளவில், கட்டின்மை எண்ணிக்கை 2ற்கான மதிப்பு 5.99. எனவே, கணிக்கப்பட்ட χ^2 மதிப்பு (4.84) ஆமோதிக்கப்பட வேண்டிய பகுதியிலேயே உள்ளது. அதனால், இல்லெனும் எடுகோளை ஆமோதிக்க வேண்டியுள்ளது. இந்த முடிவு தவறாகப் போவதற்கு 5 சதவீத வாய்ப்பே உள்ளது; 95 சதவீதம் இந்த முடிவு சரியாக இருக்கும். இதன் மூலம், மூன்று தேர்வாளர்களும் கொடுத்துள்ள தேர்ச்சி விகிதங்களுக்கு இடையே, புள்ளியியல் முக்கியத்துவம் பெறும் அளவுக்கு, வித்தியாசம் இல்லை.

பெறப்பட்ட அலைவெண்களிலிருந்து தேறாதவர்கள் விகிதங்களை கணக்கிட்டால், அன்புக்கு (5 ÷ 55) 9 சதவீதமும் ஆதிக்கு (14 ÷ 61) 23 சதவீதமும், இசைக்கு (8 ÷ 64) 12.5சதவீதமும் வருகின்றன. ஆனாலும், இவற்றிற்கிடையே உள்ள வித்தியாசங்கள், புள்ளியியல் முக்கியத்துவம் பெறாதவைகளாக உள்ளன என்று χ^2 சோதனையை வைத்து

முடிவு செய்யப்படுகிறது. இப்போது காணப்படுகின்ற வித்தியாசங்கள், வேண்டுமென்றே செய்த செயல்களினால் இருக்க முடியாது; வேறேதும் தெரியாத காரணங்களினால் விளைந்த வித்தியாசங்களாக இருக்கலாம்.

χ^2 பரவலின் மூலம் ஓர் ஈருறுப்புப் பரவல் சரியாக உள்ளதா என்பதைக் கணிக்கும் முறை கீழே தரப்படுகிறது.

நான்கு குழந்தைகள் உள்ள 160 குடும்பங்களில் கீழ்க்காணும் (அட்டவணை 55) அலைவெண் பரவல் கிடைத்தது. இப்பரவல் ஈருறுப்புக் கோட்பாட்டுக்குப் பொருத்தமாக உள்ளதா எனப் பார்க்கலாம்.

அட்டவணை - 55
சேகரிக்கப்பட்ட புள்ளிவிபரங்கள்

ஆண்/பெண் குழந்தைகளின் எண்ணிக்கை	4 ஆண் குழந்தைகள் 0 பெண் குழந்தை	3ஆண் குழந்தைகள் 1 பெண் குழந்தை	2 ஆண் குழந்தைகள் 2 பெண் குழந்தைகள்	1 ஆண் குழந்தை 3 பெண் குழந்தைகள்	0 ஆண் குழந்தை 4 பெண் குழந்தைகள்
குடும்பங்களின் எண்ணிக்கை	9	28	55	44	24

pயை ஆண்குழந்தை பிறப்பதற்கான நிகழ்தகவாகவும் $(\frac{1}{2})$

qயை பெண்குழந்தை பிறப்பதற்கான நிகழ்தகவாகவும் $(\frac{1}{2})$

எடுத்துக்கொண்டு 4 ஆண் குழந்தைகள் முதல் 0 ஆண்குழந்தை வரை இருப்பதற்கான நிகழ்தகவைக் கணக்கிடலாம்.

$$p(4 \text{ ஆண்கள்}) = 4_{c_4}p^4q^0 = 1(\frac{1}{2})^4 = \frac{1}{16}$$

$$p(3 \text{ ஆண்கள்}) = 4_{c_3}p^3q^1 = 4(\frac{1}{2})^4 = \frac{4}{16}$$

$$p(2 \text{ ஆண்கள்}) = 4_{c_2}p^2q^2 = 6(\frac{1}{2})^4 = \frac{6}{16}$$

$$p(1 \text{ ஆண்}) = 4_{c_1} p^1 q^3 = 4\left(\frac{1}{2}\right)^4 = \frac{4}{16}$$

$$p(0 \text{ ஆண்}) = 4_{c_0} p^0 q^4 = 1\left(\frac{1}{2}\right)^4 = \frac{1}{16}$$

இந்த நிகழ்தகவுகளுடன் மொத்தக் குடும்பங்களின் எண்ணிக்கையைப் பெருக்கினால் ஈருறுப்புப்பரவல் கோட்பாட்டு அலைவெண்கள் கிடைக்கும்.

அவை முறேயே 10, 40, 60, 40, 10 ஆகும். நம் கேள்வி, சேகரிக்கப்பட்ட அலைவெண்களுக்கும் கோட்பாட்டுப்படியான அலைவெண்களுக்கும் இடையில் வித்தியாசம் இருக்கிறதா இல்லையா என்பதுவே.

இல்லெனும் எடுகோள் H_0 : அவை இரண்டுக்கும் வித்தியாசம் இல்லை. அவை இரண்டும் சமமே.

H_1 : அவை இரண்டுக்கும் இடையே வித்தியாசம் உள்ளது. எனவே கொடுக்கப்பட்டுள்ள அலைவெண் பரவல் ஈருறுப்புக் கோட்பாட்டு அலைவெண் பரவலுடன் ஒத்துப் போகவில்லை.

$$\chi^2 = \frac{(9-10)^2}{10} + \frac{(28-40)^2}{40} + \ldots + \frac{(24-10)^2}{10} =$$

$$= 0.1 + 3.6 + 0.42 + 0.4 + 19.6 = 24.12$$

95 சதவீத நம்பிக்கை அளவின்படி, 5 சதவீத முக்கியத்துவத்தின்படி, கட்டின்மை எண்ணிக்கை 4க்கு (5-1), χ^2 பட்டியல் கொடுக்கும் மதிப்பு 9.49. கணக்கிடப்பட்ட χ^2ன் மதிப்பு 24.12ஆக (அதிகமாக) இருப்பதால், இல்லெனும் எடுகோள் மறுக்கப்படுகிறது. அதாவது, கிடைத்துள்ள அலைவெண்களுக்குச் சமமாக இல்லை. எனவே, ஆண் குழந்தை, பெண் குழந்தை பிறப்பதற்கான நிகழ்தகவு $\frac{1}{2}$

என்பது சரியில்லாமல் இருக்கலாம். அல்லது, கிடைத்துள்ள அலைவெண்கள் சமவாய்ப்பு முறைப்படி எடுக்கப்படாத மாதிரிகளிலிருந்து கிடைத்தவைகளாக இருக்கலாம். அல்லது கருக்கொலை / சிசுக்கொலை அங்கு நடந்திருக்கலாம்.

நேர்வு சார்புக்கெழு (Coefficient of Contigency)

ஒரு நேர்வுப்பட்டியலில் பிரிக்கப்பட்டுள்ள (ஆண், பெண், தேறியவர்கள், தேறாதவர்கள் போன்ற) பிரிவுகளுக்கு இடையேயுள்ள உறவின் நிலையை அளப்பதற்கும் χ^2ன் மதிப்பு பயன்படுகிறது.

நேர்வு சார்புக்கெழு $= c = \sqrt{\dfrac{\chi^2}{\chi^2 + N}}$

(Coefficient of contigency)

சார்புக்கெழு அதிகமாக அதிகமாக கொடுக்கப்பட்டுள்ள பிரிவுகளுக்கிடையேயான உறவு (relationship, association, dependence) அதிகம், நெருக்கம் என்று பொருள்.

ஓர் உதாரணம் எடுத்துப் பார்க்கலாம். ஓர் ஆய்வின்போது அட்டவணை 56 கிடைத்தது.

அட்டவணை - 56
கிடைத்த அலைவெண்கள்

	குணமானவர்களின் எண்ணிக்கை	குணமாகாதவர்களின் எண்ணிக்கை	மொத்தம்
மருந்து சாப்பிட்டவர்களின் எண்ணிக்கை	75	25	100
மருந்து சாப்பிடாதவர்களின் எண்ணிக்கை	65	35	100
மொத்தம்	140	60	200

அட்டவணை - 57
எதிர்பார்க்கப்படும் அலைவெண்கள்

	குணமானவர்களின் எண்ணிக்கை	குணமாகாதவர்களின் எண்ணிக்கை	மொத்தம்
மருந்து சாப்பிட்டவர்களின் எண்ணிக்கை	70	30	100
மருந்து சாப்பிடாதவர்களின் எண்ணிக்கை	70	30	100
மொத்தம்	140	60	200

$$\chi^2 = \frac{(75-70)^2}{70} + \ldots + \frac{(35-30)^2}{30} = 2.38$$

$$C = \sqrt{\frac{2.38}{2.38+200}} = 0.1084$$

இந்த நேர்வு சார்புக்கெழு மருந்து சாப்பிட்டதற்கும் குணமடைந்ததற்கும் ஒரு சிறிய அளவு உறவுள்ளதென்று காட்டுகிறது.

அதற்குப் பதிலாக அட்டவணை 58இல் உள்ளதுபோல் அலைவெண்கள் இருந்தால், மருந்து சாப்பிடுவதற்கும் குணமடைவதற்கும் அதிக உறவு உள்ளதாக நேர்வு சார்புக்கெழு பெரியதாக இருக்கும்.

அட்டவணை - 58

	குணமானவர்களின் எண்ணிக்கை	குணமாகாதவர்களின் எண்ணிக்கை	மொத்தம்
மருந்து சாப்பிட்டவர்களின் எண்ணிக்கை	100	0	100
மருந்து சாப்பிடாதவர்களின் எண்ணிக்கை	0	100	100
மொத்தம்	100	100	200

அட்டவணை 58இல் உள்ள விபரங்கள் மருந்து சாப்பிட்ட அனைவரும் குணமடைந்ததாகவும் மருந்து சாப்பிடாதவர்கள் அனைவரும் குணமடையாததாகவும் காட்டுகிறது. எனவே, மருந்து சாப்பிடுவதற்கும் குணமடைவதற்கும் தொடர்பு இருப்பதாகவும், அவை ஒன்றை ஒன்று சார்ந்திருப்பதாகவும் தெரிகிறது.

இதற்கு χ^2 கண்டுபிடிக்கலாம்.

$$\chi^2 = \frac{(100-50)^2}{50} + \frac{(0-50)^2}{50} + \ldots + \frac{(100-50)^2}{50} = 200$$

χ^2 இங்கு Nக்குச் சமமாக இருப்பதைக் கவனிக்கலாம்.

$$C = \sqrt{\frac{200}{200+200}} = 0.7071$$

அட்டவணை 58இல் கொடுக்கப்பட்டுள்ள அலைவெண்களை மாற்றி அட்டவணை 59இல் உள்ளதுபோல் கொடுத்தால் என்ன நடக்கிறதென்று பார்க்கலாம்.

அட்டவணை - 59

கிடைத்த அலைவெண்கள்

	குணமானவர்களின் எண்ணிக்கை	குணமாகாதவர்களின் எண்ணிக்கை	மொத்தம்
மருந்து சாப்பிட்டவர்களின் எண்ணிக்கை	0	100	100
மருந்து சாப்பிடாதவர்களின் எண்ணிக்கை	100	0	100
மொத்தம்	100	100	200

$$\chi^2 = \frac{(0-50)^2}{50} + \frac{(100-50)^2}{50} + \frac{(100-50)^2}{50} + \frac{(0-50)^2}{50} = 200$$

$$C = \sqrt{\frac{200}{200+200}} = 0.7071$$

இப்பொழுது மருந்து சாப்பிடுவதற்கும், குணமாவதற்கும் தொடர்பு உள்ளது. மருந்து சாப்பிடுவதும் குணமாவதும் ஒன்றையொன்று சார்ந்த நிகழ்ச்சிகளே. ஆனால், இதற்குப் பதிலாக அலைவெண்கள் அட்டவணை 60இல் உள்ளது போல் இருந்தால், மருந்து சாப்பிடுவதும் குணமாவதும் ஒன்றையொன்று சாராத நிகழ்வுகளாகத் தோன்றும். ஏனெனில், அங்கு கிடைத்த அலைவெண்களும், எதிர்பார்க்கப்படும் அலைவெண்களும் சமமாக இருக்கும். எனவே $\chi^2 = 0$. எனவே $c = 0$.

அட்டவணை - 60

	குணமானவர்களின் எண்ணிக்கை	குணமாகாதவர்களின் எண்ணிக்கை	மொத்தம்
மருந்து சாப்பிட்டவர்களின் எண்ணிக்கை	50	50	100
மருந்து சாப்பிடாதவர்களின் எண்ணிக்கை	50	50	100
மொத்தம்	100	100	200

அட்டவணை 60க்கு எதிர்பார்க்கப்படும் அலைவெண்கள் உட்கட்டங்கள் நான்கிலுமே 50 தான் வரும்.

எனவே $$\chi^2 = \Sigma \frac{(o - e)^2}{e} = 0$$

$$c = \sqrt{\frac{0}{0 + 200}} = 0$$

பண்புகளின் ஒட்டுறவு

χ^2யைப் பயன்படுத்தி பண்புகளின் ஒட்டுறவுக்கெழுவும் கண்டுபிடிக்கலாம். உதாரணமாக, அட்டவணை 56 மற்றும் 57களில் காணப்படும் விபரங்களுக்கு,

$$r = \sqrt{\frac{\chi^2}{n\,(k-1)}} = \sqrt{\frac{2.38}{200}} = 0.1091$$

இந்த r எடுத்துக்கொண்ட இரண்டு குணங்களுக்கிடையே மிக மிகக் குறைவான உறவு இருப்பதாகக் காட்டுகிறது. அட்டவணை 58இல் காணப்படும் புள்ளி விபரங்களுக்கு,

$$r = \sqrt{\frac{200}{200\,(k-1)}} = \sqrt{\frac{200}{200}} = 1$$

அட்டவணை 58இல் உள்ள விபரங்களின்படி, எடுத்துக் கொண்ட இரு குணங்கள் (மருந்து சாப்பிடுவது, குணமடைவது) பூரண நேரிடை உறவைக் கொண்டுள்ளது என்பதை r காட்டுகிறது.

F பரவல் (F DISTRIBUTION)

இதுவரை இயல்நிலை, ஈருறுப்பு, t, பாய்ஸான், பல்லுறுப்பு, அதிபெருக்கு மற்றும் χ^2 ஆகிய பரவல்கள் பற்றி விவரிக்கப்பட்டன. இவையனைத்தும் முழுமையின் சில பண்பலகுகளைப் (parameters) பற்றி மதிப்பீடு செய்வதற்கும், அவற்றிற்கிடையே உள்ள உறவுகளையும் வித்தியாசங்களையும் ஒப்பிட்டுப் பார்ப்பதற்கும் பயன்படுத்தப்படுவதாக அறியப்பட்டது. ஒட்டுறவுக்கெழுக்கள் மற்றும் உடன்தொடர்புப் போக்குக்கெழுக்களின் புள்ளியியல் முக்கியத்துவத்தைச் சோதித்துப் பார்க்கவும் z, t பரவல்கள் பயன்படுமென விளக்கப்பட்டது.

இரண்டுக்கும் மேற்பட்ட முழுமைகளையோ அவற்றின் பண்பலகுகளான கூட்டுச் சராசரிகளையோ அல்லது இரண்டு மாறுபாடுகளுக்கு (variances) இடையேயுள்ள வித்தியாசங்களையோ அளந்து பார்த்து அவற்றின் புள்ளியியல் முக்கியத்துவத்தைப் புரிந்து கொள்ள F பரவல் உள்ளது. இந்த F பரவல் பற்றிய கருத்துக்களை 1920ன் ஆரம்ப காலங்களில் ஆர்.எ.ஃபிஷர் (R.A.FISHER) என்பவர் அறிமுகப்படுத்தினார். எனவே, அவருக்கு மரியாதை செய்யும் பொருட்டு, 'F' பரவலெனப் பெயரிடப்பட்டது.

இந்த F பரவல் இரண்டு மாறுபாடுகளை ஒப்பிட்டுப் பார்ப்பதற்கும், இரண்டுக்கும் மேற்பட்ட மாதிரிகளின் கூட்டுச் சராசரியை ஒப்பிட்டுப் பார்ப்பதற்கும், பல மாறிகளுக்கு இடையேயுள்ள ஒட்டுறவுக்கெழுவின் புள்ளியியல் முக்கியத்துவத்தை அறியச் செய்யவும் பயன்படுகிறது. F பரவலில் F மதிப்பு கணக்கிடும்போது மாறுபாடுகள் வித்தியாசமாக இருந்தால், அதிகமான மாறுபாட்டை பகுதியாகவும் (numerator) குறைவான மாறுபாட்டை தொகுதியாகவும் கொள்வதால், Fன் மதிப்பு ஒன்று அல்லது ஒன்றுக்கு மேல்தான் எப்போதும் இருக்கும். Fன் அதிகபட்ச மதிப்பை அளவிடமுடியாது.

உதாரணமாக, மாறுபாடுகளின் அளவினை முதலில் ஒப்பிடலாம். இருவேறு விதமான (அ, ஆ) முறைகளில் தயாரிக்கப்பட்ட விளக்குகளின் திறன்கள் இயல்நிலைப் பரவலை ஒத்திருக்கின்றன. 'அ' முறையில் தயாரிக்கப்பட்ட 17 விளக்குகள் சமவாய்ப்பு முறையில் எடுக்கப்பட்டபோது அவற்றின் திறனின் திட்டவிலக்கம் 60 மணிநேரமாக இருந்தது. இது 'ஆ' முறையில் தயாரிக்கப்பட்ட 21 விளக்குகளுக்கு 50 மணி நேரமாக இருந்தது. இந்த இரு மாறுபாடுகளுக்கும் உள்ள வித்தியாசம் சந்தர்ப்பவசத்தால் வந்ததா? அல்லது, இருமுறைகளுக்கும் இடையே உண்மையிலேயே முக்கியமான வேறுபாடுகள் உள்ளனவா?

இதற்கு, முதலில் இல்லெனும் எடுகோள் கொள்ள வேண்டும்.

$$H_0 : \sigma_1 = \sigma_2$$

மாற்று எடுகோள் : $H_1 : \sigma_1 \neq \sigma_2$

$$F_c = \frac{\hat{\sigma}^2_{அ}}{\hat{\sigma}^2_{ஆ}} = \frac{n_{அ} s^2_{அ}}{n-1} \div \frac{n_{ஆ} s^2_{ஆ}}{n-1}$$

$$= \frac{(17)\,3600}{16} \div \frac{21\,(2500)}{20}$$

$$= \frac{61200}{16} \div \frac{52500}{20}$$

$$= 3825 \div 2625$$

$$= 1.46$$

F பரவலின் பட்டியல் மதிப்பு *(16, 20, 5%) 2.18.*

அதாவது, F, 2.18யைவிட அதிகமாக இருப்பதற்கான நிகழ்தகவு *0.05*. கணக்கிடப்பட்ட F ஒத்துக் கொள்ளும் எல்லைக்குள் வருவதால், இல்லெனும் எடுகோளை ஆமோதிக்கலாம். அப்படியானால், இரண்டு முறைகளின் திட்டவிலக்கங்களுக்கு இடையே உள்ள வேறுபாடு சந்தர்ப்பவசத்தால் நிகழ்ந்ததென முடிவு செய்யலாம்.

விகிதங்களில் உள்ள வேறுபாடுகளைச் சோதனை செய்யவும் (Taro Yamane, Statistics : An Introductory Analysis, Third Edition, p. 812) 'F' பரவல் பயன்படுத்தப்படுகிறது.

உதாரணமாக, ஒரு போட்டித் தேர்வில் 5 சதவீத மாணவர்கள் தேர்ச்சி பெற்றுள்ளார்கள். ஒரு நூறு கூறுகளைக் கொண்ட மாதிரி எடுத்தபோது அதில் மூன்று மாணவர்கள் தேர்ச்சி பெற்றிருந்தனர். அப்படியானால் தேர்ச்சி விகிதம் குறைவாக உள்ளது என முடிவு செய்யலாமா? இதற்கு இல்லெனும் எடுகோள்:

$H_0 : P = 5\%$

$H_1 : P < 5\%$

$$\phi_1 = 2\,(K+1) = 2(3+1) = 8$$

$$\phi_2 = 2\,(n\text{-}K) = 2(100\text{-}3) = 194$$

$$F = \frac{\phi_2 P}{\phi_1 Q} = \frac{(194)(0.05)}{(8)(0.95)} = \frac{9.7}{7.6} = 1.28$$

5 சதவீத முக்கியத்தவத்தில் $F^{8}_{194} > 1.98$. எனவே, இல்லெனும் எடுகோள் ஒத்துக் கொள்ளப்படுகிறது. அதாவது முழுமையில் உள்ள தேர்ச்சி விகிதத்திற்கும், மாதிரியில் உள்ள தேர்ச்சி விகிதத்திற்கும் உள்ள வித்தியாசம் புள்ளியியல் முக்கியத்துவம் பெறவில்லை. எனவே, அங்கு காணப்படும் வித்தியாசம் சந்தர்ப்பவசமானதே.

இதேகணக்கினை இயல்நிலைப் பரவலைக் கொண்டும் செய்து பார்க்கலாம். அப்படிச் செய்யும்போது $z_c = 0.91$ வரும். இதுவும், இயல்நிலைப் பரவலின் பட்டியல் மதிப்பைவிடச் சிறியதானதே. எனவே, இல்லெனும் எடுகோள் ஒத்துக் கொள்ளப்படும். காணப்படும் வித்தியாசம் சந்தர்ப்பவசமானதே என முடிவு செய்யப்படும்.

$$z_c = \frac{p - P}{\sqrt{\frac{PQ}{n}}} = \frac{-0.02}{\sqrt{\frac{0.05 \times 0.95}{100}}} = -0.91$$

பரவற்படி ஆய்வு (Analysis of Variance) அல்லது மாறுபாடு ஆய்வு

மாறுபாடு ஆய்வு ஆர்.எ.ஃபிஷர்-ஆல் உருவமைக்கப் பட்டது. இம்முறை முதலில் விவசாய ஆய்வுகளில் பயன்படுத்தப்பட்டது. பின்னர் மற்ற துறைகளிலும் இதன் பயன்பாடு பரவியது (R.A.Fisher, The Design of Experiments, 4th Edition, Edinburgh : Oliver and Boyd, 1947). இப்போது இதன் உபயோகம் அதிகரித்துக் கொண்டுவருகிறது. (உ.ம். W.G.Cochran, G.M.Cox, O.Kempthrone, W.T. Federer, G.R.Snedecor ஆகியவர்களின் ஆய்வுகள்)

இரண்டு கூட்டுச் சராசரிகளை ஒப்பிட்டுப் பார்க்க z, t பயன்படுகிறது. ஒவ்வொரு சமயமும் இரண்டு இரண்டு

கூட்டுச்சராசரிகளாக எடுத்து (உ.ம். $\mu_1\mu_2$; $\mu_1\mu_3$; $\mu_2\mu_3$) z மற்றும் t மூலம் ஆய்வு செய்யலாம். ஆனால் அவ்வாறு செய்வதால் நேரம் அதிகம் விரயமாகலாம். இதைத் தவிர்க்கவே, மாறுபாடு ஆய்வு பயன்படுத்தப்படுகிறது. இதில் கீழ்வரும் பகுதிகள் உள்ளன.

1. அனைத்து வர்க்கங்களின் கூடுதல்
   (Total Sum of Squares = TSS)

2. குழுக்களுக்குள்ளேயான வர்க்கங்களின் கூடுதல்
   (Sum of Squares Within groups = WSS)

3. குழுக்களுக்கிடையேயான வர்க்கங்களின் கூடுதல்
   (Sum of Squares Between groups = BSS)

உடன் தொடர்பு ஆய்வில் (Regression Analysis), குழுக்களுக்கிடையேயான வர்க்கங்களின் கூடுதல், விளக்கப்பட்ட வேறுபாடுகள் (explained variations) என்றும், குழுக்களுக்குள்ளேயான வர்க்கங்களின் கூடுதல், விளக்கப்படாத வேறுபாடுகள் (unexplained variations) என்றும் அழைக்கப்படுகின்றன. இவ்வாறாக, உடன்தொடர்பு ஆய்வும் மாறுபாடு ஆய்வும் நெருங்கிய தொடர்புள்ளனவாக இருக்கின்றன.

இப்பொழுது மூன்று குழுக்களின் கூட்டுச் சராசரிகளுக்கு இடையே உள்ள வித்தியாசம் புள்ளியியல் முக்கியத்துவம் வாய்ந்ததா அல்லது மாதிரிகள் எடுப்பில் உள்ள பிழைகளால் வந்ததா என்று பார்க்கலாம். இதனைச் செய்வதற்கு வெவ்வேறு புத்தகங்கள் வெவ்வேறு முறைகளைக் கடைப்பிடிக்கின்றன. பல்வேறு முறைகளில், ஜான் டபிள்யூ.பெஸ்ட் மற்றும் ஜேம்ஸ் வி.கான் (John W. Best & James V. Kahn, 2006, Research in Education, Ninth Edition, Prentice Hall of India, New Delhi, 2006) கடைப்பிடித்துள்ள (பக்கம் 410-412) முறை சற்று மாற்றி இங்கு தரப்படுகிறது.

அட்டவணை - 61

மாணவிகள்	1	2	3	4	5	6	7	8	9	10	மொத்தம்
கணிதத்தில் மதிப்பெண்கள்	18	22	18	23	19	24	20	21	19	25	209
பொருளியலில் மதிப்பெண்கள்	26	27	18	22	23	19	27	26	24	26	238
தமிழில் மதிப்பெண்கள்	18	14	15	14	19	21	17	17	18	19	172

முதலில் எடுகோள்கள் தயார் செய்து கொள்ள வேண்டும்.

இல்லெனும் எடுகோள் : $H_0 : \mu_1 = \mu_2 = \mu_3$

மாற்று எடுகோள் : $H_0 : \mu_1 \neq \mu_2 \neq \mu_3$

$$F = \frac{\text{குழுக்களுக்கிடையேயான வர்க்கங்களின் கூட்டுச்சராசரி}}{\text{குழுக்களுக்குள்ளேயான வர்க்கங்களின் கூட்டுச்சராசரி}}$$

அதாவது : $\frac{MS_b}{MS_w}$

$$MS_b = \frac{SS_b}{dof}$$

$$SS_b = \frac{\Sigma (X_1)^2}{n_1} + \frac{\Sigma (X_2)^2}{n_2} + \dots + \frac{\Sigma (X_n)^2}{n_n} - CF$$

dof = கட்டின்மை எண்ணிக்கை

$$CF = \frac{(GT)^2}{N}; \; MSw = \frac{SS_w}{dof}; \; SS_w = SS_t - SS_b$$

SS_t சில புத்தகங்களில் TSS (Total Sum Square) என்றும் குறிக்கப்பட்டுள்ளது.

கொடுக்கப்பட்டுள்ள ஒவ்வொரு Xக்கும் வர்க்கம் கண்டுபிடித்து, அவற்றைக் கூட்டி, வந்த தொகையிலிருந்து CFயைக் கழிக்க வேண்டும்.

முதலில் CF (திருத்த ம் காரணி - Correction Factor)யைக் கண்டுபிடிக்கலாம்.

$$CF = \frac{(GT)^2}{N}$$

GT = Grand Total = 209 + 238 + 172 = 619

N = 30

$$\boxed{CF} = \frac{(619)^2}{30} = \boxed{12772.03}$$

$$\boxed{SS_t} = (18)^2 + (22)^2 + ... + (26)^2 + (27)^2 + ... + (18)^2 + (19)^2 - CF$$

$$= 13{,}191 - 12772.03 = \boxed{418.97}$$

$$\boxed{SS_b} = \frac{(209)^2}{10} + \frac{(238)^2}{10} + \frac{(172)^2}{10} - 12{,}772.03$$

$$= 218.87$$

$$\boxed{SS_w} = SS_t - SS_b$$

$$= 418.97 - 218.87 = \boxed{200.10}$$

Total Sum Square = Between Sum Square + Within Sum Square

என்றும் சில புத்தகங்களில் குறிப்பிடப்படுகிறது.

அதாவது,(1) அனைத்து எண்களின் வர்க்கங்களின் கூடுதல் =

(2) குழுக்களுக்கிடையேயான வர்க்கங்களின் கூடுதல் +

(3) குழுக்களுக்குள்ளேயான வர்க்கங்களின் கூடுதல்

இதிலிருந்து இது உடன் தொடர்பு ஆய்வுடன் நெருங்கிய தொடர்பு கொண்டுள்ளதெனத் தெரிகிறது. மேலே கொடுக்கப்பட்டுள்ளவற்றில் முதலாவது மொத்த

வேறுபாடுகளையும் (total variation) இரண்டாவது விளக்கப்பட்ட வேறுபாடுகளையும் (explained variation) மூன்றாவது விளக்கப்படாத வேறுபாடுகளையும் (unexplained variation) காட்டுகின்றன. கொடுக்கப்பட்டுள்ள உதாரணத்தில்

$$\boxed{MS_b} = \frac{218.87}{2} = \boxed{109.44}$$

இதில் கட்டின்மை எண்ணிக்கை 2, ஏனெனில் மொத்தம் 3 குழுக்கள் உள்ளன (d.o.f. = n - 1 = 3 - 1 = 2).

$$\boxed{MS_w} = \frac{200.1}{27} = \boxed{7.41}$$

இதில் கட்டின்மையின் எண்ணிக்கை 27. ஏனெனில் ஒவ்வொரு குழுவிலும் *10* மாணவர்கள் இருப்பதாலும், அங்கு *3* குழுக்கள் இருப்பதாலும்

(10-1) + (10-1) + (10-1) = 9 + 9 + 9 = 27.

மூன்று குழுக்கள் கொண்ட இந்தப் பரவற்படி ஆய்வின் சுருக்கம் அட்டவணை 62இல் கொடுக்கப்படுகிறது.

அட்டவணை - 62

மாறுபாடுகளுக்கான மூலம்	வர்க்கங்களின் கூடுதல்	கட்டின்மை எண்ணிக்கை	கூட்டுச் சராசரி	F
குழுக்களுக்கிடையே (பாடங்கள்)	218.87	2	109.44	14.77
குழுக்களுக்குள்ளே (பிழை)	200.10	27	7.41	--
மொத்தம்	418.97	--	--	--

F பட்டியலில் *5* சதவீத முக்கியத்துவத்திற்கு *95* சதவீத நம்பிக்கை அளவுக்கு, பொருத்தமான கட்டின்மை

எண்ணிக்கைகளுக்குத் (2,27) தேவையான மதிப்பைப் பார்த்தால், அது 3.34க்கும் 3.37க்கும் இடையில் இருக்கிறது. ஆனால், கணிக்கப்பட்ட மதிப்பு (F_c) அதைவிட அதிகமாகவே உள்ளது. 1 சதவீத முக்கியத்துவத்திற்குப் பார்த்தால் கூட பட்டியல் மதிப்பு F_T, 5.45க்கும் 5.53க்கும் இடையில் உள்ளது. எனவே கொடுக்கப்பட்டுள்ள இல்லெனும் எடுகோளை 99 சதவீத நம்பிக்கையுடன் நிராகரிக்கலாம். அல்லது, அந்த மூன்று கூட்டுச் சராசரிகளும் சமமில்லை என்று 99 சதவீத நம்பிக்கையுடன் சொல்லலாம்.

இதுவரை நாம் பார்த்தது ஒரு வழி பரவற்படி ஆய்வாகும். (One Way ANOVA). கொடுக்கப்பட்டுள்ள எடுத்துக்காட்டில் பாடங்களுக்குள்ளே வேறுபாடுகள் இருந்தனவா என்று பார்த்தோம். மாணவர்களுக்குள்ளே வேறுபாடுகள் இருந்தனவா என்று பார்க்கவில்லை. மாணவர்களுக்குள்ளேயும் வேறுபாடுகள் இருக்கலாம். அப்படிப் பார்த்தால், அதற்கு இருவழி பரவற்படி ஆய்வு என்று பெயர்.

இருவழிப் பரவற்படி ஆய்வுக்கு ஓர் உதாரணம் எடுக்கலாம். நான்கு வகையான நெல்வகைகள் உள்ளன. அவற்றின் பெயர் A_1, A_2, A_3 மற்றும் A_4. இந்த நான்கு நெல்வகைகளும் மூன்று விதமான நிலங்களில் (B_1, B_2, B_3) பயிரிடப்படுகின்றன. நிலம் ஒவ்வொன்றும் 1 ஏக்கர் பரப்பளவு கொண்டது. விளைச்சல் அளவு டன்னில் கொடுக்கப்பட்டுள்ளது (அட்டவணை 63). விளைச்சலில் நிலவகைகளுக்கு இடையே வேறுபாடு உள்ளதா என்றும், நெல்வகைகளுக்கு இடையே வேறுபாடு உள்ளதா என்றும் கண்டுபிடிக்க வேண்டும்.

அட்டவணை - 63

	B_1	B_2	B_3	Total
A_1	2	3	4	9
A_2	4	6	8	18
A_3	7	7	10	24
A_4	8	9	9	26
Total	**21**	**25**	**31**	**77**

அட்டவணை - 64

அட்டவணை 63இல் உள்ள எல்லா எண்களின் வர்க்கங்கள்

	B_1	B_2	B_3	Total
A_1	4	9	16	81
A_2	16	36	64	324
A_3	49	49	100	576
A_4	64	81	81	676
Total	**441**	**625**	**961**	**5929**

குறிப்பு : இந்த அட்டவணை (64)இல் கடைசி நிரலும் நிரையும் உள்ளிருக்கும் எண்களின் மொத்தம் அல்ல.

திருத்தக்காரணி $\boxed{CF} = \frac{(77)^2}{12} = \boxed{494.08}$

$\boxed{TSS}$ (Total Sum Square)

(அனைத்து எண்களின் வர்க்கங்களின் கூடுதல்) =

$= 2^2 + 4^2 + \ldots + (10)^2 + (9)^2 - CF$

$= 569 - 494.08 = \boxed{74.92}$

RSS (Row Sum Square)

(நிரைகளின் மொத்தங்களின் வர்க்கங்களின் கூடுதல்) =

$$= \frac{81}{3} + \frac{324}{3} + \frac{576}{3} + \frac{676}{3} - 494.08 =$$

$$= 27 + 108 + 192 + 225.33 - 494.08 = \boxed{58.25}$$

CSS (Column Sum Square)

(நிரல்களின் மொத்தங்களின் வர்க்கங்களின் கூடுதல்) =

$$= \frac{441}{4} + \frac{625}{4} + \frac{961}{4} - CF =$$

$$= 110.25 + 156.25 + 240.25 - 494.08$$

$$= \boxed{12.67}$$

ESS (Error Sum Square)

(பிழை வர்க்கக் கூடுதல்)

= TSS - [RSS + CSS]

$$= 74.92 - [58.25 + 12.67] = \boxed{4}$$

அட்டவணை - 65

மூலம்	வர்க்கங்களின் கூடுதல்	கட்டின்மை எண்ணிக்கை	கூட்டுச் சராசரி	F
நிரை	58.25	4-1 = 3	19.40	$\frac{19.4}{0.67} = 29.0$
நிரல்	12.67	3-1 = 2	6.34	$\frac{6.34}{0.67} = 9.46$
பிழை	4.00	3 × 2 = 6	0.67	
மொத்தம்	74.92	12-1 = 11		

F பட்டியலின் மதிப்புக்கள்

$F_6^3\ (0.05) = 4.76$

$F_6^3\ (0.01) = 9.78$

$F_6^2\ (0.05) = 5.14$

$F_6^2\ (0.01) = 10.90$

கணிக்கப்பட்டுள்ள F மதிப்புக்கள் (29.0, 9.46) F பட்டியல் மதிப்புக்களுடன் ஒப்பிடப்படும்போது, நிரைகளுக்குள்ளும் நிரல்களுக்குள்ளும் உள்ள வேறுபாடுகள் புள்ளியியல் முக்கியத்துவம் பெற்றவை என்று 95 சதவீத நம்பிக்கையுடன் சொல்ல முடியும். இதிலிருந்து, நிலங்களுக்கிடையேயும் விளைச்சலில் வித்தியாசம் உள்ளது. நெல்வகைகளுக்குள்ளும் விளைச்சலில் வித்தியாசம் உள்ளது. இந்த வித்தியாசங்கள் சந்தர்ப்பவசத்தால் நிகழ்ந்தவை அல்ல.

மேலே கூறப்பட்டவைகளைத் தவிர, தற்போது வரும் கணினிச் சிப்பங்கள், F பரவலை வைத்து பல மாறி ஒட்டுறவுக் கெழுக்களைச் சோதனை செய்யவும் வழி செய்துள்ளன.

மேலும், பரவற்படி ஆய்வில் இதுவரை ஒரு மாறி (பாடங்கள் - முதல் உதாரணம்) இருமாறி (இரண்டாவது உதாரணம் - நெல்வகைகள், நில வகைகள்) வேறுபாடுகளைச் சோதனை செய்வது எப்படி என்று விளக்கப்பட்டது. இதனுடைய அடுத்தகட்டமாக, பல மாறிகளுக்கு (MANOVA, Multiway ANOVA) இடையே உள்ள வேறுபாடுகளை ஆய்வு செய்யவும் வழிவகைகள் செய்யப்பட்டுள்ளன.

பண்பலகில்லாச் சோதனைகள் (NONPARAMETRIC TESTS)

இதுவரையில், மாதிரிகளின் குணங்களைக் கொண்டு (Statistics உ.ம். சராசரிகள், மாறுபாடுகள்) முழுமை அல்லது முழுமைகளின் பண்பலகுகளைச் (Parameters) சோதனை செய்வது பற்றி விளக்கப்பட்டது. பல சமயங்களில்,

பரவல்களுக்குள் கொண்டுவர முடியாத இயல்புகளைப் பற்றியும் பண்பலகுகள் இல்லாத புள்ளி விபரங்களையும் ஆராய வேண்டிய சூழ்நிலை உருவாகிறது.

கீழ்வரும் சூழ்நிலைகள் பண்பளவைகள் அல்லாத சோதனைகளுக்குப் பொருத்தமாக அமைகின்றன.

1. மாதிரிகள் எடுக்கப்படும் முழுமைகள் இயல்நிலைப் பரவலாக இல்லாமல் இருக்கலாம்.
2. மாறிகள் பகுதிகளாகப் பிரிக்கப்பட்டு அலைவெண்கள் தரப்பட்டிருக்கலாம் (nominal forms, classified in categories)
3. மாறிகள் தரப்படுத்தப்பட்டவைகளாக (ordinal form, ranked in order) இருக்கலாம்.

பண்பலகு மற்றும் பண்பலகில்லாச் சோதனைகளின் முக்கியத்துவங்கள் பற்றி வெவ்வேறு கருத்துக்கள் நிலவுகின்றன. பண்பலகுச் சோதனைகளின் வெற்றி முழுமையின் பரவல் முறையைப் பொறுத்து அமைகிறது. ஆனால், அவற்றிற்குச் சிறிதும் முக்கியத்துவம் கொடுக்காமல் பண்பலகுச் சோதனைகள் பல சமயங்களில் தவறாக நடத்தப்படுகின்றன. இதுபோல் தவறுதலாகச் சோதனைகள் செய்வதைவிட, சரியானப் பரவலுக்குள் வராத முழுமைகளுக்கு, நேரடியாகப் பண்பலகில்லாச் சோதனைகளைப் பயன்படுத்துவது சிறந்தது என்கின்றனர் சிலர்.

பல பண்பலகில்லாச் சோதனைகள் உள்ளன. அவற்றில் χ^2 சோதனையும் மேன்-விட்னி (MANN-WHITNEY) சோதனையும் மிகவும் அதிகமாகப் பயன்படுத்தப்படுகின்றன என்கிறார்கள் பெஸ்டும் கானும் (John W. Best & James V.Kahn, Research Education, Prentice - Hall of India, Ninth Edition, 2006, p. 419)

பண்பலகில்லாச் சோதனைகளில் முதலாவதாக χ^2 சோதனை பல புத்தகங்களில் தரப்படுகின்றது. அதில், இரண்டு அல்லது அதற்கு மேற்பட்ட மாறிகள் சார்ந்துள்ளனவா இல்லையா எனக் கண்டுபிடிப்பது முதல்வகை பண்பலகில்லாச் சோதனையாக உள்ளது. இது பற்றி முன்னரே பார்க்கப்பட்டது. இருப்பினும் தொடர்ச்சியாக இருப்பதற்காக இன்னுமொரு உதாரணம் இங்கு தரப்படுகிறது. புகைப்பிடித்தலுக்கும் படிப்பறிவுக்கும் தொடர்பு உள்ளதா? அவை ஒன்றை ஒன்று சார்ந்துள்ளனவா? சாராதவைகளா? அவை சார்ந்திருப்பது போலத் தோன்றினால் உண்மையிலேயே அப்படி உள்ளனவா? அல்லது அந்த உறவு சந்தர்ப்பவசத்தால் அல்லது மாதிரிப் பிழை (Sampling Error)யால் அப்படித் தெரிகிறதா போன்ற கேள்விகளுக்கு χ^2 சோதனை மூலம் பதில் சொல்லலாம்.

அட்டவணை - 66

	புகைப்பவர் எண்ணிக்கை	புகைக்காதவர் எண்ணிக்கை	மொத்தம்
படிப்பறிவு பெற்றவர் எண்ணிக்கை	40	10	50
படிப்பறிவு பெறாதவர் எண்ணிக்கை	5	45	50
மொத்தம்	**45**	**55**	**100**

முதலில் இல்லெனும் எடுகோள் உருவாக்கப்பட வேண்டும்.

H_0 : படிப்பறிவுக்கும் புகைப்பதற்கும் எந்தவித உறவும் இல்லை. அவையிரண்டும் சார்பிலா நிகழ்வுகளே.

H_1 : படிப்பறிவும் புகைப்பிடித்தலும் ஒன்றையொன்று சார்ந்துள்ளன.

$$\chi^2 = \Sigma \frac{(O-E)^2}{E}$$

இரண்டாவதாக, எதிர்பார்க்கப்படும் அலைவெண்களைக் கணக்கிட்டு அவற்றைப் பயன்படுத்தி χ^2 கணக்கிடவேண்டும். (அட்டவணை 67).

அட்டவணை - 67

E		O	$\frac{(O-E)^2}{E}$
$E_{11} = \frac{50 \times 45}{100} =$	22.50	40	13.61
$E_{12} = \frac{50 \times 55}{100} =$	27.50	10	11.14
$E_{21} = \frac{50 \times 45}{100} =$	22.50	5	13.61
$E_{22} = \frac{50 \times 55}{100} =$	27.50	45	11.14
மொத்தம்		**100**	**49.50**

$\chi^2_c = 49.50$. χ^2 பட்டியல் மதிப்பு {5 சதவீத முக்கியத்துவத்திற்கு கட்டின்மை எண்ணிக்கை [(2-1) (2-1)] ஒன்றுக்கு} 3.84

χ_c^2 யின் மதிப்பு χ_t^2யின் மதிப்பைவிட அதிகமாக இருப்பதால், இல்லெனும் எடுகோள் நிராகரிக்கப்படுகிறது. எனவே, படிப்பறிவும் புகைப்பிடித்தலும் ஒன்றையொன்று சார்ந்தே உள்ளன. இதை 99 சதவீத நம்பிக்கையோடு கூடச் சொல்லலாம். ஏனெனில் 1 சதவீத முக்கியத்துவத்திற்கு χ^2ன் மதிப்பு 6.64, கணிக்கப்பட்ட χ^2ன் மதிப்பைவிடக் குறைவு தான். எனவே, நாம் சொல்லும் இந்த முடிவு தவறாகப் போவதற்கான வாய்ப்பு மிக மிகக் குறைவே. அதாவது, தவறாகப் போவதற்கான நிகழ்தகவு 0.01க்கும் குறைவே.

மேன்-விட்னி சோதனை (MANN - WHITNEY TEST)

கிடைத்துள்ள மதிப்புக்கள் ஏதேனும் ஒரு வரிசையில் (order) அடுக்கப்பட்டு இருந்தால் இந்தச் சோதனை பயன்படுத்தப்படுகிறது. கிடைத்துள்ள புள்ளி விபரங்களின்

எண்ணிக்கை 20க்கும் மேல் சென்றால், இயல்நிலைப் பரவலைப் போன்ற முடிவினையே இந்தச் சோதனையும் தரும்.

மேன்-விட்னி சோதனையைப் புரிந்து கொள்ள ஜான் டபிள்யூ பெஸ்ட் மற்றும் ஜேம்ஸ் வி.கான் ஆகியவர்கள் எழுதியுள்ள புத்தகத்தில் (Research in Education) 425ஆம் பக்கத்தில் எடுத்துள்ள உதாரணத்தை எடுத்துக் கொள்ளலாம். ஒரு நல்ல ஆசிரியர் தன் வகுப்பில் உள்ள 20 மாணவர்களுக்கு, இரண்டு விதமாகப் பாடம் நடத்தி மதிப்பெண்கள் அளித்துள்ளார். 20 மாணவர்களும் முதல் முறையில் புரிந்து தேர்வில் பெற்ற மதிப்பெண்களும் இரண்டாம் முறையில் புரிந்து தேர்வில் பெற்ற மதிப்பெண்களும் அட்டவணை 68இல் கொடுக்கப்பட்டுள்ளன. இந்த இரண்டு முறைகளுக்கிடையே வித்தியாசம் உள்ளதா? இருந்தால், அது சந்தர்ப்பவசமாகவோ மாதிரிப் பிழையினாலோ வந்ததா? அல்லது வித்தியாசம் புள்ளியியல் முக்கியத்துவம் பெற்றதா? என்று அறிய வேண்டும்.

அட்டவணை - 68

முதல் முறையினால் பெற்ற மதிப்பெண்கள்	வ.எண்.	2ஆம் முறையினால் பெற்ற மதிப்பெண்கள்	வ.எண்.
50	3	47	1
52	5	49	2
60	8	51	4
63	10	55	6
68	13	56	7
72	17	61	9
75	20	64	11
77	22	66	12

78	23	69	14
79	24	70	15
80	25.5	71	16
80	25.5	73	18
82	28	74	19
83	29	76	21
85	31	81	27
88	33.5	84	30
89	35	87	32
94	38	88	33.5
95	39	90	36
97	40	92	37

இதற்கு, இல்லெனும் எடுகோள் (H_0) : இரண்டு முறைகளினால் மாணவர்கள் பெற்ற மதிப்பெண்களுக்கு இடையே வித்தியாசம் இல்லை. இதைச் சோதனை செய்ய மேன்-விட்னி சோதனை பயன்படுத்தப்படுகிறது. இச்சோதனைக்குத் தேவையானவை u_1, u_2 மற்றும் z

$$u_1 = N_1N_2 + \frac{N_1(N_1+1)}{2} - \Sigma R_1$$

$$= (20)(20) + \frac{20(21)}{2} - 469.50 = 140.50$$

N_1, N மாணவர்களின் எண்ணிக்கை முதல் முறையிலும் இரண்டாவது முறையிலும்

$\Sigma^{r}{}_1$ என்பது முதல் முறையில் பங்குபெற்றுள்ள மாணவர்களின் வரிசை எண்களின் கூடுதல், வரிசையெண்கள் மதிப்பெண்களின் அடிப்படையில் ஏறுமுகமாக அளிக்கப்பட்டுள்ளது. இந்தக் கணக்கில் இது 469.50 முதல் வகைக்கு.

$$u_2 = N_1N_2 + \frac{N_2(N_2+1)}{2} - \Sigma R_2$$

$$= (20)(20) + \frac{20(21)}{2} - 350.50 = 259.50$$

u_2 யைப் பெற இன்னொரு வழியும் உண்டு.

அதாவது $u_1 = N_1N_2 - u_2$

$$140.50 = 400 - u_2; \quad u_2 = 400 - 140.50 = 259.50$$

அடுத்ததாக, கிடைத்துள்ள மதிப்புக்களைப் பயன்படுத்தி z கண்டுபிடிக்க வேண்டும்.

$$z_c = \frac{u_1 - \frac{N_1N_2}{2}}{\sqrt{\frac{N_1N_2(N_1+N_2+1)}{12}}} = \frac{140.50 - \frac{400}{2}}{\sqrt{\frac{400(41)}{12}}} = -1.61$$

z பட்டியலில் 5 சதவீத முக்கியத்துவத்திற்கான மதிப்பு 1.96. இங்கு கணிக்கப்பட்ட zன் மதிப்பு ஆமோதிக்கும் பரப்புளவுக்கு உள்ளே இருக்கிறது. எனவே, இல்லெனும் எடுகோள் ஆமோதிக்கப்படுகிறது. அதாவது, மாணவர்களின் திறனைக் கூட்டுவதில் இரண்டு முறைகளுக்கும் இடையே வித்தியாசம் இல்லை. முதல் முறையில் மாணவர்கள் அதிக மதிப்பெண்கள் பெற்று உயர்ந்த வரிசை எண்களைப் பெற்று வரிசை எண்களின் கூடுதலை அதிகமாகப் பெற்றிருந்த போதும், இது ஒரு சந்தர்ப்பவசமான நிகழ்வேயொழிய முதல் முறை புள்ளியியல் முக்கியத்துவத்தில் சிறந்த முறை என்று சொல்ல முடியாது. இந்த முடிவை 95 சதவீத நம்பிக்கையுடன் சொல்லலாம். இந்த முடிவு தவறாவதற்கு வாய்ப்பு மிகக்குறைவே; அதாவது இம்முடிவு தவறாவதற்கான நிகழ்தகவு 0.05 தான்.

பண்பளவையிலாச் சோதனைகள் இன்னும் பல உள்ளன. ஒரு மாதிரி குறிச்சோதனை (One-sample sign test) இரு-மாதிரி குறிச்சோதனை (two-sample sign test) இருமாதிரி இடைநிலைச்

சோதனை (two-sample Median test) இரண்டுக்கும் அதிகமான மாதிரி இடைநிலைச் சோதனை (K-sample Median test) வில்காக்ஸன் பொருத்தப்பட்ட இணை சோதனை (WILCOXON MATCHED-PAIRS TEST) க்ரஸ்கல் - வால்லிஸ் சேர்தனை (KRUSKAL - WALLIS TEST) ஒரு மாதிரி ஓட்டச் சேர்தனை (One-sample runs test) தால்மோகோரோவ் - ஸ்மிர்னவ் ஒரு மாதிரி சோனை (KOLMOGOROV - SMIRNOV ONE - SAMPLE TEST) ஆகிய பண்பளவைச் சோதனைகளுக்கு விளக்கங்கள் ஜி.சி.பெரி எழுதியுள்ள (Business Statistics, Third Edition, Tata McGraw Hill Education Private Limited, New Delhi, 2010) புத்தகத்தில் 599ஆவது பக்கம் முதல் 632ஆம் பக்கம் வரை கிடைக்கின்றன.

ஓர் எச்சரிக்கை (A word of caution)

பலவகையான சோதனைகள் (tests) இருந்தபோதும், அனுமானங்கள் இல்லாத முழுமையிலிருந்து எடுக்கப்படும் மாதிரிகளை வைத்துச் செய்கின்ற சோதனைகளின் திறனும் சிறப்பும் குறைவே. எனவே அனுமானங்கள் கொண்ட முழுமையைப் பயன்படுத்தும் சோதனைகளே அதிக அளவு நம்பிக்கைக்குரியவை. அனுமானங்களைக் கொண்ட முழுமையை வைத்து நடத்தும் சோதனைகளின் நம்பிக்கை அளவை எட்ட, சிறிய அளவு மாதிரியே போதும். ஆனால் அதே அளவு நம்பிக்கையை அடைய அனுமானங்கள் இல்லாத முழுமைகளிலிருந்து பெரிய அளவு மாதிரியை எடுக்க வேண்டியிருக்கும். அதிக அனுமானங்களுடன், அதிக செய்திகளை ஒரு மாதிரியிலிருந்து பெற முடியும். ஆனால் அதே சமயம் அதிக அனுமானங்கள் செயல்முறைகளின் பயன்பாட்டினைக் குறைக்கலாம். இந்த எச்சரிக்கையினை ஃப்ரூண்டு மற்றும் வில்லியம்ஸ் (JOHN E.FREUND & FRANK J. WILLIAMS, Elementary Business Statistics : The Modern Approach, Prentice Hall Inc., New Jersey, 1982) தங்கள் புத்தகத்தில் 426ஆம் பக்கத்தில் குறிப்பிட்டுள்ளார்கள்.

௸

10. காலம்சார் தொடர்வரிசை ஆய்வு (TIME SERIES ANALYSIS)

புள்ளி விபரங்களைப் பல வகைகளாகப் பார்க்கிறோம். அவற்றில் ஒருவகை காலத்தைக் கருத்தில் கொண்டு பிரிப்பது. புள்ளி விபரங்கள் ஒரு குறிப்பிட்ட நேரத்தில் சேகரிக்கப்படலாம்; ஒரு குறிப்பிட்ட நேரத்துக்கான புள்ளிவிபரங்கள் சேகரிக்கப்படலாம். இந்த நேரம் பகல் 12 மணியாகவோ மாலை 5 மணியாகவோ, ஒரு நாளாகவோ, ஒரு மாதமாகவோ, ஒரு வருடமாகவோ, ஓரைந்து ஆண்டுகளாகவோ ஒரு பத்தாண்டுகளாகவோ இருக்கலாம். இப்படி ஏதேனும் ஒரு குறித்த நேரத்திற்கு, காலத்திற்கு பல நபர்கள், வீடுகள், தெருக்கள், கிராமங்கள், வட்டங்கள், மாவட்டங்கள், மாநிலங்கள், நாடுகள் பற்றிய புள்ளி விபரங்கள் சேகரிக்கப்பட்டு இருந்தால் அவை குறுக்குவெட்டுப் புள்ளிவிபரங்கள் (cross-section data) என அழைக்கப்படுகின்றன. உதாரணமாக 2009ஆம் ஆண்டு மாநில வாரியாக எவ்வளவு பெண்சிசுக்கொலை நடந்திருக்கிறது என்று புள்ளிவிபரம் கிடைத்தால் அவை குறுக்குவெட்டுப் புள்ளிவிபரங்கள் ஆகும். ஒரு குறிப்பிட்ட நாள், ஓர் ஊரில் உள்ள ஒவ்வொரு வீட்டிலும் எவ்வளவு பணம் மருந்துக்காக செலவு செய்துள்ளார்கள் என்பது குறுக்கு வெட்டுப் புள்ளிவிபரம் ஆகும். மாவட்டம் வாரியாக 2009ஆம் ஆண்டு எத்தனை விவசாயிகள் தற்கொலை செய்துள்ளனர் என்பது குறுக்குவெட்டுப்புள்ளி விபரம் ஆகும். மாறாக இந்தியாவில் ஒவ்வொரு மாதமும் எத்தனை பெண் கருக்கொலைகள் நடந்துள்ளன என்பது காலம்சார் புள்ளி விபரமாகும். தமிழகத்தில் ஒவ்வொரு ஆண்டும் எத்தனை சாலை விபத்துக்கள் நடந்துள்ளன என்பது காலம்சார் புள்ளி விபரமாகும்.

ஒரே ஒரு பொருளைப் பற்றி அல்லது உறவைப் பற்றி ஆய்வு செய்யும்போது இருவேறு வகையான புள்ளி விபரங்கள் வெவ்வேறு முடிவுகளைத் தரலாம். வருமானம் கூடும்போது, நுகர்வுக்காகும் செலவின் விகிதம் குறைந்து கொண்டு செல்லும் (நுகர்வுக்காகும் மொத்தச் செலவு கூடிக்கொண்டு செல்லலாம்) என்று ஒருவர் சொல்லலாம். இது ஒரு குறிப்பிட்ட காலத்தில் பல வீடுகள் / நபர்கள் இடமிருந்து சேகரிக்கப்பட்ட குறுக்குவெட்டுப் புள்ளி விபரங்கள் தரும் செய்தியாக இருக்கலாம். ஆனால், ஒரு நாடு ஒவ்வொரு ஆண்டும் நுகர்வுக்காகச் செலவு செய்த விகிதம் (மொத்த வருமானத்தில்) கூடிக்கொண்டு சென்றுள்ளது என்று மற்றவர் சொல்லலாம். இது காலம்சார் புள்ளிவிபரத்திலிருந்து கிடைக்கின்ற செய்தியாக இருக்கலாம். இவ்வாறாக, ஒரே பொருள் பற்றிய ஆய்வின் முடிவை வெவ்வேறாகக் காலம்சார் தொடர்வரிசைப் புள்ளி விபரங்களும், குறுக்குவெட்டுப் புள்ளிவிபரங்களும் தரலாம். எனவே, ஓர் ஆய்வின் முடிவைப் பற்றி பேசும்போது, அந்த ஆய்வில் எப்படிப்பட்ட புள்ளிவிபரங்கள் பயன்படுத்தப்பட்டுள்ளன என்று பார்ப்பது அவசியமாகிறது.

புள்ளி விபரங்களில் பல சார்புள்ள மாறிகளாக இருக்கின்றன. ஒவ்வொரு வகையான புள்ளிவிபரமும் மற்றொரு வகையான புள்ளி விபரத்துடன் தொடர்புள்ளதாக இருக்கிறது. ஒன்று மற்றொன்றைப் பாதிக்கவும் செய்யலாம்; ஒன்று மற்றொன்றால் பாதிக்கப்படவும் செய்யலாம்.

பல சூழ்நிலைகளில் காலம் ஒரு முக்கியமான காரணியாகி விடுகிறது. அதனால்தானோ என்னவோ, சிலர் அடிக்கடி ‘இதெல்லாம் என் நேரம்’ என்கிறார்கள். நேரம் நிச்சயம் பல நடவடிக்கைகளை நிர்ணயம் செய்கின்றது. சூரிய வெப்பத்தின் அளவு, குளிரின் அளவு, மழையின் அளவுபோல இன்னும் பல காலத்தால் மாறுபடுகின்றன.

'காலம் கருதுதல்' என்பது மிக முக்கியம் என்று வள்ளுவரும் கூறியுள்ளார். 'ஆடிப்பட்டம் தேடி விதை' என்றும் சொல்வார்கள். 'காலத்தினால் செய்த உதவி' 'காற்றுள்ள போதே தூற்றிக்கொள்' என்பதெல்லாம் அன்றாடம் கேட்கும் அர்த்தமுள்ள சொற்றொடர்கள். காலம் மாறுவதாலும், காலத்திற்கேற்ப சூழ்நிலைகள் மாறுவதாலும், மனிதர்களின் நடவடிக்கைகளும் மாறுகின்றன. எனவே, சமூகப் பொருளாதார அரசியல் நிகழ்ச்சிகளும் காலத்தால் மாற்றியமைக்கப்படுகின்றன. ஆழமாக யோசித்துப் பார்த்தால் எல்லாமே நேரம்தான் என்று கூடச் சொல்லலாம்.

சமகால இடைவெளியில் ஒவ்வொரு காலத்திலும் (மணி, நாள், வாரம், மாதம், வருடம்) ஒரு மாறி பெறும் மதிப்புகளை வரிசையாக அமைத்துப் பெறும் தொடரை காலம்சார் தொடர் வரிசை என்கிறார்கள்.

உதாரணமாக, கடந்த சில ஆண்டுகளாக குழந்தைகளின் பாலின விகிதம் (Child Sex Ratio) பெண் குழந்தைகளுக்கு எதிராகக் குறைந்து கொண்டே செல்கிறது. இது செழிப்பான மாநிலங்கள் என்று இந்தியாவில் சொல்லப்படும் கேரளா, பஞ்சாப் ஆகிய மாநிலங்களிலும் உண்மை. இது ஒரு காலம்சார் தொடர்வரிசை எனலாம்.

பலவகைப்பட்ட நிகழ்ச்சிகளால் காலத்தொடர் வரிசையில் மாற்றங்கள் ஏற்படுகின்றன. இவற்றில் சில குறிப்பிட்ட கால இடைவெளிவிட்டு அடுத்தடுத்து நிகழும் தன்மை வாய்ந்தவை. இதில் கால இடைவெளியென்பது மாறலாம். மாதமாக இருக்கலாம், பருவமாக இருக்கலாம், ஆண்டாக இருக்கலாம், பல ஆண்டுகள் சேர்ந்த ஒரு நீண்ட காலமாக இருக்கலாம்.

காலத்தொடர் வரிசையில் ஏற்படும் மாற்றங்களை நான்கு பகுதிகளாகப் பிரிக்கலாம்.

1. நீண்ட காலப் போக்கு (Seculaı Trend) → T

2. பருவகால மாற்றங்கள் (Seasonal variations) → S

3. சுழல் ஏற்ற இறக்கங்கள் (Cyclical fluctuations) → C

4. ஒழுங்கற்ற ஏற்ற இறக்கங்கள் (Irregular or erratic fluctuations) → I

நீண்டகாலப்போக்கு எந்தவிதமாகவும் இருக்கலாம். நேர்கோடுகளாகவோ வளைகோடுகளாகவோ இருக்கலாம். உதாரணமாக, மக்கள்தொகை வளர்ச்சி என்பது முதலில் குறைவான வேகத்தில் வளர்ந்தும் பின்னர் அதிகமான வேகத்தில் வளர்ந்தும் இருக்கலாம்.

காலத்தொடர் வரிசை முறைகள் (Time Series Models)

காலத்தொடர் வரிசையானது அதன் நான்கு பகுதிகளையும் கூட்டினால் கிடைப்பதாக மேற்கொண்டால்,

$Y = T + S + C + I \Rightarrow$ (Additional Model) கூட்டல் முறை.

இதில் Y = கொடுக்கப் பெற்றுள்ள காலத் தொடர் வரிசையிலுள்ள மதிப்புக்கள்.

சில சமயங்களில், கொடுக்கப் பெற்றுள்ள தொடர்வரிசையின் மதிப்புக்கள் நான்கு பகுதிகளின் பெருக்குத் தொகைக்குச் சமம் என மேற்கொள்ளப்படலாம். அப்போது,

$Y = T \times S \times C \times I$ ஆக இருக்கும்.

இதைப் பெருக்கல் முறை (Multiplicative model) என்கிறார்கள்.

நீண்ட காலம் என்பது எல்லா இடங்களிலும் பல வருடங்களைக் குறிக்க வேண்டிய அவசியம் இல்லை. கொடுக்கப்பட்டுள்ள விபரங்களின் ஒரே சீரான மாற்றத்தை அறிவதற்குப் போதுமான அளவு நீளமாக உள்ள காலத்தையே நீண்டகாலம் என குறிக்கிறார்கள். எனவே நீண்டகாலம் என்பது

ஒரு நாளோ, ஒரு மாதமோ, ஓர் ஆண்டோ, பல ஆண்டுகளாகவோ இருக்கலாம். இது எடுக்கின்ற பொருளைப் பொறுத்தது. உதாரணமாக, பாக்டீரியா உற்பத்தியைப் பொறுத்தவரை சில நாட்களையே நீண்ட காலமாகக் கருதலாம். ஏனெனில், பாக்டீரியா உற்பத்தியை 5 நிமிடங்களுக்கு ஒருமுறை எண்ணி, சில நாட்களிலேயே அதன் உற்பத்திப்போக்கை அறிந்துவிடலாம். நீண்ட ஆண்டுகள் காத்திருக்க வேண்டிய அவசியம் இல்லை. ஆனால் அதே சமயம், ஓர் இரும்பு உற்பத்தி செய்யும் ஆலையின் உற்பத்திப்போக்கைச் சில நாட்களில் அறிய முடியாது; பல ஆண்டுகள் தேவைப்படலாம். எனவே, இங்கு பல ஆண்டுகள் சேர்ந்ததே நீண்ட காலம் ஆகும்.

நீண்டகாலப் போக்கு என்பது மேல்நோக்கித்தான் இருக்கவேண்டும் என்ற அவசியம் இல்லை; எப்படி வேண்டுமானாலும் இருக்கலாம்.

நீண்டகாலப் போக்கினை கீழ்க்காணும் ஏதேனும் ஒரு முறையில் அளக்கலாம்.

1. தடையின்றி கையினால் வரையும் முறை (Free hand method)

2. நகரும் சராசரி முறை (Moving average method)

3. அரைச் சராசரி முறை (Semi-average method)

4. குறைந்த வர்க்க முறை (Method of least squares)

பருவகால மாற்றங்கள் (Seasonal variations)

பருவகாலம் என்பது ஓர் ஆண்டிற்குள் குறைந்தகால இடைவெளியில் விட்டுவிட்டு நிகழும் காலத்தைக் குறிக்கிறது. பொதுவாக, பருவகால மாற்றங்கள் ஒரு வாரம் அல்லது ஒரு மாதம் அல்லது மூன்று மாதங்கள் இடைவெளிவிட்டு நிகழும் தன்மையுடையவை. தட்பவெப்ப நிலையில் ஏற்படும் மாற்றங்கள் பருவகால மாற்றங்களைப்

பெரிதும் பாதிக்கிறன. இம்மாற்றங்கள் வேளாண் பொருள் உற்பத்தியையும் அதன் மூலமாக மற்ற தொழில்களின் உற்பத்தியையும், மக்களின் தேவையையும் பெரிதும் பாதிக்கின்றன.

சுழல் ஏற்ற இறக்கங்கள் (Cyclical variations)

இவையும் ஒழுங்காக நிகழும் தன்மை வாய்ந்தவை. ஆனால், ஓர் ஆண்டிற்கு மேற்பட்ட கால இடைவெளியில் விட்டுவிட்டுத் திரும்பத் திரும்ப நிகழும் தன்மை உடையவை. இவ்வித ஏற்ற இறக்கங்கள் 5 முதல் 10 ஆண்டுகளுக்குள் நிகழலாம். இந்தக்கால இடைவெளி ஒவ்வொரு சுழலின் போதும் சமமாக இல்லாமல் இருக்கலாம்.

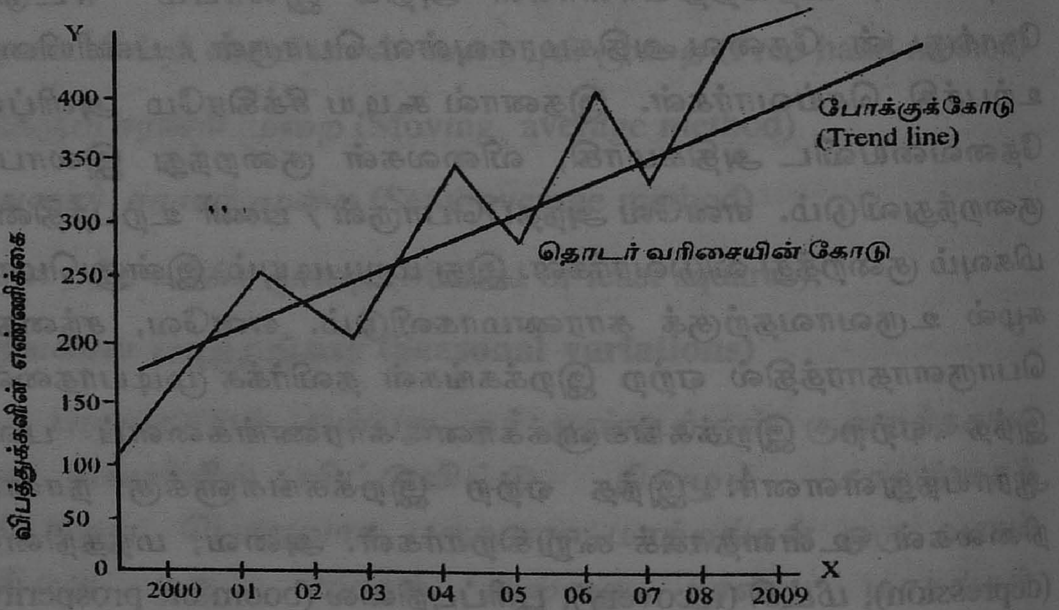
சுழல் ஏற்ற இறக்கங்களை 'வாணிபச் சுழல்கள்' (Business Cycles or Trade Cycles) என்று கூறுவதுண்டு. முதலாளித்துவ பொருளாதார அமைப்பில் ஏற்ற இறக்கங்கள் தவிர்க்க முடியாதவையாகும். சந்தையில் ஏதேனும் ஒரு பொருள் அல்லது பணிக்கு அதிகத் தேவை வந்து, அவற்றின் விலை கூடினால், உற்பத்தியாளர்கள் அதிக இலாபம் ஈட்டும் நோக்குடன் தேவை அதிகமாகவுள்ள பொருள் / பணியினை உற்பத்தி செய்வார்கள். இதனால் கூடிய சீக்கிரமே அளிப்பு, தேவையைவிட அதிகமாகி, விலைகள் குறைந்து இலாபம் குறைந்துவிடும். எனவே அந்தப்பொருள் / பணி உற்பத்தியை மிகவும் குறைத்து விடுவார்கள். இது மறுபடியும் இன்னுமொரு சுழல் உருவாவதற்குக் காரணமாகவிடும். எனவே, சந்தைப் பொருளாதாரத்தில் ஏற்ற இறக்கங்கள் தவிர்க்க முடியாதவை. இந்த ஏற்ற இறக்கங்களுக்கான காரணங்களைப் பலர் ஆராய்ந்துள்ளனர். இந்த ஏற்ற இறக்கங்களுக்கு நான்கு நிலைகள் உள்ளதாகக் கூறுகிறார்கள். அவை, மந்தநிலை (depression), மீட்சி (recovery), பூரிப்புநிலை (boom or prosperity) மற்றும் பின்னிறக்கம் (recession or decline) ஆகியவை ஆகும்.

சந்தைப் பொருளாதாரம் ஊகவாணிகத்திற்கு (speculation) வழிவகுப்பதால், பல பொருளாதாரக் குற்றங்கள் (பதுக்கல் போன்றவை) நிகழ வாய்ப்புள்ளது. செல்வந்தர்கள் அரசை மிரட்டிப் பணம் சம்பாதித்து, வாணிபச் சுழல்களைத் தங்களுக்கு பலன் தரும் வகையில் மாற்றிக் கொள்ளும் (2008-2009இல் நடந்ததுபோல) திறன் படைத்தவர்களாக இருப்பதால், மீண்டும் மீண்டும் ஏற்ற இறக்கங்களை உருவாக்கிக் கொண்டிருப்பார்கள். இதுவே, பிறகு பழகிப் போய்விடும்.

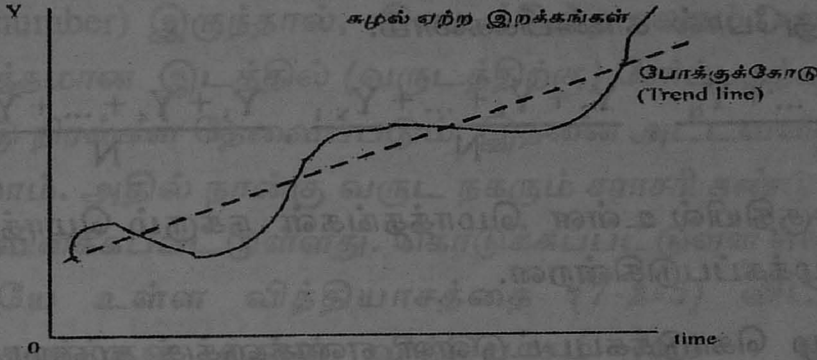
ஒழுங்கற்ற ஏற்ற இறக்கங்கள் (Irregular and erratic fluctuations)

மனிதர்களின் கொள்கைகளினால் உருவாக்கப்படும் வன்முறைச் செயல்களும், இயற்கைச் சீற்றங்களும் இதற்கு உதாரணங்கள். இந்நிகழ்வுகள் நிகழ்வதற்குக் குறிப்பிட்ட காலம் ஏதும் கூறமுடியாது. சுனாமி (Tsunami) போன்றவை யாராலும் கண்டுபிடிக்கப்பட முடியாமலேயே பல நாடுகளைத் தாக்கும் திறன் படைத்தவையாக உள்ளன.

வரைபடம் - 39



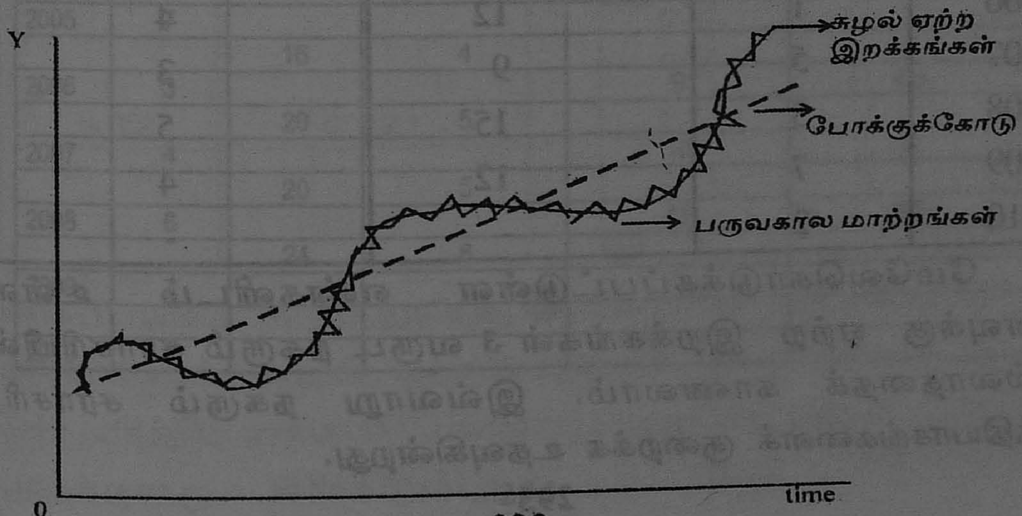
வரைபடம் - 40



தடையின்றி கையினால் வரையும் முறை (Freehand method)

பொதுவாக சார்பற்ற, தன்னிச்சையான (independent) மாறியை X அச்சிலும் சார்ந்திருக்கும் மாறியை Y அச்சிலும் கொள்வது வழக்கமாக உள்ளது. நாம் அன்றாடம் பார்க்கும் சாலை விபத்துக்களின் எண்ணிக்கை ஆண்டுக்கு ஆண்டு கூடிக்கொண்டே வருகிறது என்று கொள்வோம். அது வரைபடம் 39இல் காட்டப்படுகிறது. வரைபடம் 40 போக்குக் கோடுகளுடன் சுழல் ஏற்ற இறக்கங்களையும் காட்டுகிறது. வரைபடம் 41 போக்குக் கோடு, சுழல் ஏற்ற இறக்கங்களுடன் பருவகால மாற்றங்களையும் காட்டுகிறது.

வரைபடம் - 41



நகரும் சராசரி மூலம் (Moving average method)
நீண்டகாலப் போக்கினை அளவிடுதல்

N வரிசையைக் கொண்ட நகரும் சராசரியைக் கீழ்வருவதுபோல் காண்பிக்கலாம்.

$$\frac{Y_1 + Y_2 + \ldots + Y_N}{N}, \frac{Y_2 + Y_3 + \ldots + Y_{N+1}}{N}, \frac{Y_3 + Y_4 + \ldots + Y_{N+2}}{N} \ldots$$

இதில் பகுதியில் உள்ள மொத்தங்கள் நகரும் மொத்தங்கள் என அழைக்கப்படுகின்றன.

கீழே கொடுக்கப்பட்டுள்ள எண்களுக்கு மூன்று வரிசை (order 3) நகரும் சராசரியைக் காணலாம்.

2, 6, 1, 5, 3, 7, 2

$$\frac{2+6+1}{3}, \frac{6+1+5}{3}, \frac{1+5+3}{3}, \frac{5+3+7}{3}, \frac{3+7+2}{3}$$

இவற்றின் சராசரிகளை அதற்குப் பொருத்தமான இடங்களில் குறிப்பது சரியாகும். எனவே அட்டவணை 69 தரப்படுகிறது.

அட்டவணை - 69

ஆண்டு	எண்கள்	3வருட நகரும் மொத்தம்	3 வருட நகரும் சராசரி
2004	2	-	-
2005	6	9	3
2006	1	12	4
2007	5	9	3
2008	3	15	5
2009	7	12	4
2010	2	-	-

மேலேகொடுக்கப்பட்டுள்ள எண்களிடம் உள்ள அளவுக்கு ஏற்ற இறக்கங்கள் 3 வருட நகரும் சராசரியில் இல்லாததைக் காணலாம். இவ்வாறு நகரும் சராசரி, வித்தியாசங்களைக் குறைக்க உதவுகின்றது.

கொடுக்கப்பட்டுள்ள வரிசை (order) ஓர் ஒற்றைப்படை எண்ணாக (odd number) இருந்தால் (மேலே உள்ளது போல்) சிறிது எளிதாக இருக்கும். அதுவே இரட்டைப்படை எண்ணாக (even number) இருந்தால், கிடைக்கின்ற நகரும் சராசரியைப் பொறுத்தமான இடத்தில் (வருடத்திற்கு) சேர்ப்பதற்காகக் கூட இரண்டு நிரல்கள் தேவைப்படும். இதனை அட்டவணை 70இல் காணலாம். அதில் நான்கு வருட நகரும் சராசரி கண்டுபிடிக்கும் முறை விளக்கப்பட்டுள்ளது. கொடுக்கப்பட்டுள்ள எண்களுக்கு இடையே உள்ள வித்தியாசத்தை (7-2=5) விட நகரும் சராசரிகளுக்குள் உள்ள வித்தியாசம் (5.5 - 4.5 = 1) குறைந்திருப்பதைக் காணலாம். மேலும், கொடுக்கப்பட்டுள்ள எண்களில் உள்ள அளவுக்கு ஏற்ற இறக்கங்கள் இல்லாமல், நகரும் சராசரியில் உள்ள ஏற்ற இறக்கங்கள் குறைவாக இருப்பதையும் காணலாம். இது ஏனெனில், நகரும் சராசரி, பருவ, சுழல் மற்றும் ஒழுங்கற்ற ஏற்ற இறக்கங்களை நீக்கிவிட்டு போக்கினை (trend) மட்டும் காட்டுகிறது.

அட்டவணை - 70

ஆண்டு	மதிப்பு	நான்கு ஆண்டு நகரும் மொத்தம்	நான்கு ஆண்டு நகரும் சராசரி	நான்கு ஆண்டு நகரும் சராசரிகளின் 2 ஆண்டு மொத்தம்	மையப்படுத்தப்பட்ட நான்கு ஆண்டு நகரும் சராசரி
2004	2				
2005	7				
		16	4		
2006	3			9	4.5
		20	5		
2007	4			10	5
		20	5		
2008	6			11	5.5
		24	6		
2009	7				
2010	7				

அரைச் சராசரி முறை (Semi-average method)

மொத்தம் கொடுக்கப்பட்டுள்ள கால இடைவெளியை இரண்டாகப் பிரித்து, அவ்விரண்டு பகுதிகளுக்கும் தனித்தனியாக கூட்டுச் சராசரி கண்டுபிடிப்பது அரைச் சராசரி முறை என அழைக்கப்படுகிறது. கொடுக்கப்பட்டுள்ள வருடங்களின் எண்ணிக்கை ஒற்றைப்படை (odd) எண்ணாக இருந்தால் மையப்பகுதியில் இருக்கும் எண்ணைப் பயன்படுத்தாமல் விட்டுவிட வேண்டும். கிடைத்துள்ள அரைச் சராசரிகளைக் கொண்டு போக்குக்கோடு (trend line) ஒன்று வரையலாம். அதன் மூலம் ஏற்ற இறக்கம் இல்லாத போக்கினை அறியமுடியும். இது வரைபடம் 42இல் தரப்படுகிறது.

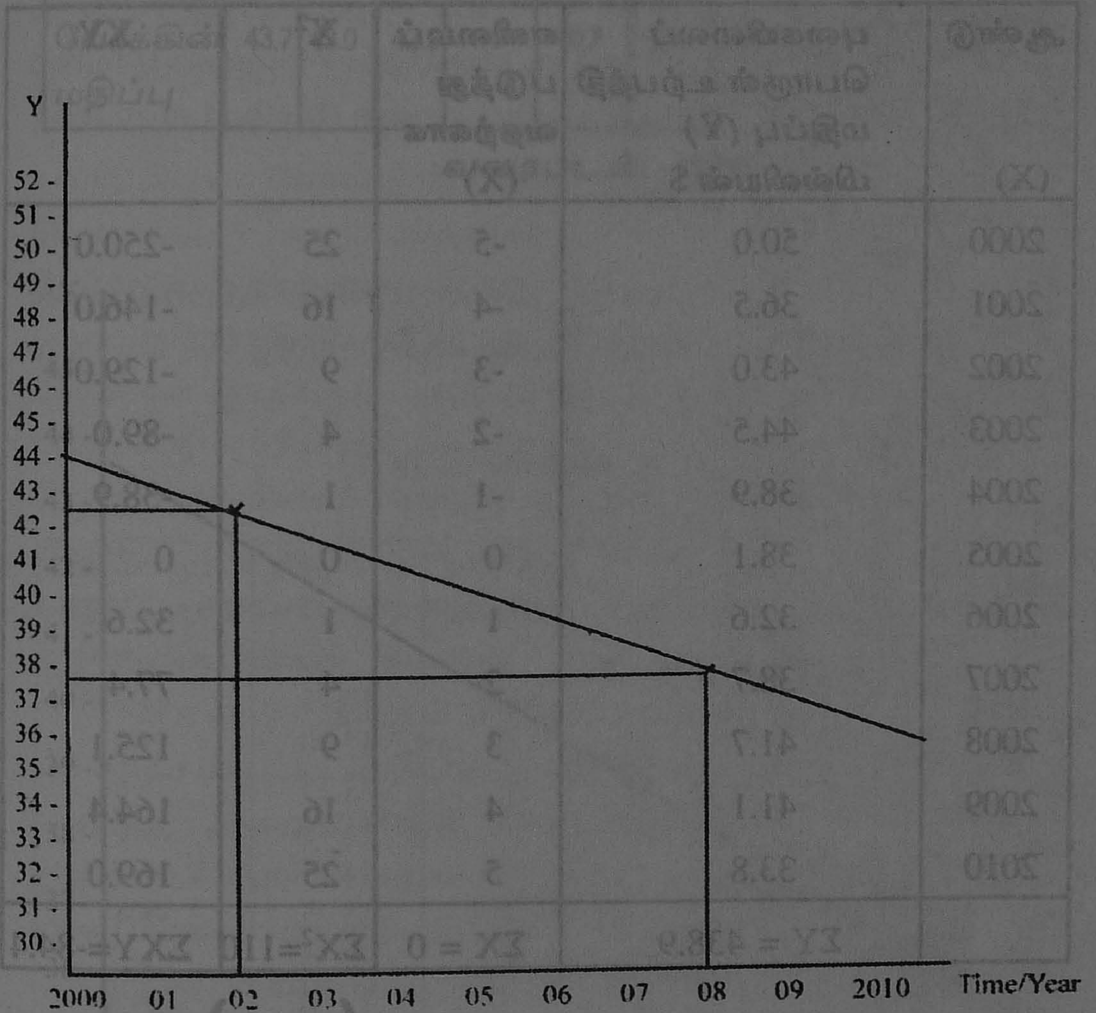
அட்டவணை - 71

ஆண்டு	புள்ளிவிபரம்	பாதியின் மொத்தம்	அரைச் சராசரி
2000	50.0		
2001	36.5		
2002	43.0	**212.9**	**42.6**
2003	44.5		
2004	38.9		
2005	38.1		
2006	32.6		
2007	38.7		
2008	41.7	**187.9**	**37.6**
2009	41.1		
2010	33.8		

2000 முதல் ஒவ்வொரு ஆண்டுக்கும் உரிய புள்ளி விபரங்களை போக்குக் கோட்டிலிருந்து பெறலாம். 2008க்கும் 2002க்கும் ஆன புள்ளிகளுக்கு இடையேயான வித்தியாசம் 42.6-37.6=5.0. இது ஆறு ஆண்டுகளில் குறைந்தது.

அப்படியானால், ஒரு ஆண்டுக்குக் குறைந்த அளவு 5÷6=0.83; இதுதான் போக்குக்கோட்டின் சாய்வு (slope அல்லது first order differential of the function). 2002இல் புள்ளி 42.6 எனில், 2003இல் இது 42.60 - 0.83 = 41.77ஆகவும் 2004இல் 41.77 - 0.83 = 40.94 ஆகவும் குறைந்து கொண்டே செல்லும்.

வரைபடம் - 42



குறைந்த வர்க்க முறை (Least square method)

மற்ற முறைகளைவிட இது மிக முக்கியமாக விளங்குவதற்குக் காரணம் இந்த முறை எல்லா புள்ளி விபரங்களுக்கும் தகுந்த முக்கியத்துவம் கொடுக்கிறது. மிகப்

பலரால் இந்த முறை பயன்படுத்தப்படுகிறது. இம்முறை மூலம் போக்குக்கோடு எவ்வாறு மதிப்பிடப்படுகிறது என்பதைக் கீழ்வரும் உதாரணம் மூலம் காணலாம். அட்டவணை 72இல் ஒரு நாடு உற்பத்தி செய்த புகையிலைப் பொருள்களின் மதிப்பு மில்லியன் $ல் தரப்பட்டுள்ளது.

அட்டவணை -72

ஆண்டு (X)	புகையிலைப் பொருள் உற்பத்தி மதிப்பு (Y) மில்லியன் $	எளிமைப் படுத்துவதற்காக (X)	X^2	XY
2000	50.0	-5	25	-250.0
2001	36.5	-4	16	-146.0
2002	43.0	-3	9	-129.0
2003	44.5	-2	4	-89.0
2004	38.9	-1	1	-38.9
2005	38.1	0	0	0
2006	32.6	1	1	32.6
2007	38.7	2	4	77.4
2008	41.7	3	9	125.1
2009	41.1	4	16	164.4
2010	33.8	5	25	169.0
	$\Sigma Y = 438.9$	$\Sigma X = 0$	$\Sigma X^2 = 110$	$\Sigma XY = -84.4$

குறைந்த வர்க்கக்கோடு $= Y = \bar{Y} + \left(\frac{\Sigma XY}{\Sigma X^2}\right) X$

$$= \frac{438.9}{11} + \left(\frac{-84.4}{110}\right) X$$

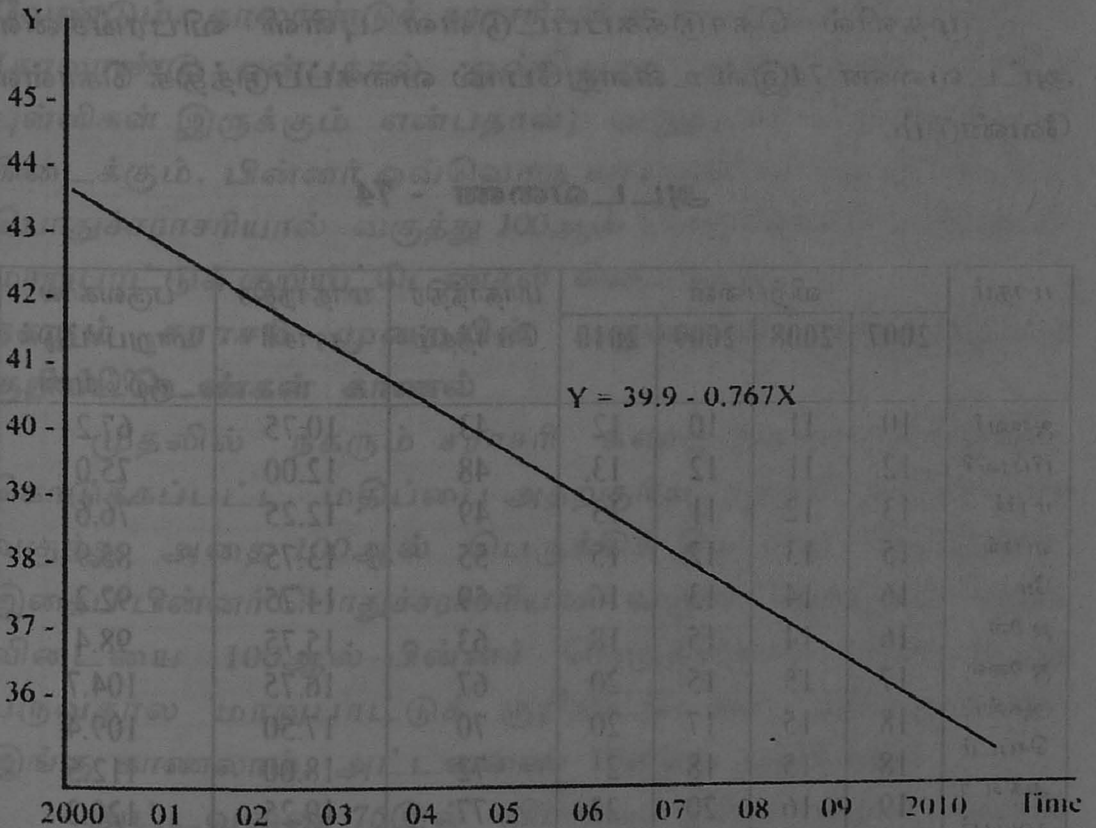
$$Y = 39.9 - 0.767 X$$

எண்களைச் சிறிய எண்களாக மாற்றுவதற்காக 2005ஆம் ஆண்டு 0 என்று கொள்ளப்பட்டது. இதனை மனதில் கொண்டு ஒவ்வொரு ஆண்டுக்கும் போக்கின் (trend) மதிப்பைக் கணக்கிட்டால் அட்டவணை 73ம் வரைபடம் 43ம் கிடைக்கும்.

அட்டவணை - 73

வருடம்	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
போக்கின் மதிப்பு	43.7	43.0	42.2	41.4	40.7	39.9	39.1	38.4	37.6	36.8	36.1

வரைபடம் - 43



பருவகால மாறுபாடுகள், சுழல் ஏற்ற இறக்கங்கள், ஒழுங்கற்ற (erratic) ஏற்ற இறக்கங்கள் ஆகியவைகளை மதிப்பீடு செய்யும் முறைகள்

காலத்தொடர் வரிசையில், போக்கினைத் (Trend) தவிர மேலேகூறப்பட்ட மூன்றுவகை மாற்றங்கள் உள்ளன. பருவகால மாறுபாடுகளை மதிப்பீடு செய்வதற்கான, நடைமுறையில் பெரிதும் பயன்படுத்தப்பெற்று வரும், முறைகளில் இரண்டு உள்ளன. *1.* எளிய சராசரி முறை (Method of simple average) (2) நகரும் சராசரி முறை (Method of moving average)

எளிய சராசரி முறை

முதலில் கொடுக்கப்பட்டுள்ள புள்ளி விபரங்களை அட்டவணை 74இல் உள்ளதுபோல் வகைப்படுத்திக் கொள்ள வேண்டும்.

அட்டவணை - 74

மாதம்	விற்பனை				மாதாந்திர மொத்தம்	மாதாந்திர சராசரி	பருவகால மாறுபாட்டுக் குறியெண்
	2007	2008	2009	2010			
ஜனவரி	10	11	10	12	43	10.75	67.2
பிப்ரவரி	12	11	12	13	48	12.00	75.0
மார்ச்	13	12	11	13	49	12.25	76.6
ஏப்ரல்	15	13	12	15	55	13.75	85.9
மே	16	14	13	16	59	14.75	92.2
ஜூன்	16	14	15	18	63	15.75	98.4
ஜூலை	17	15	15	20	67	16.75	104.7
ஆகஸ்ட்	18	15	17	20	70	17.50	109.4
செப்டம்	18	15	18	21	72	18.00	112.5
அக்டோ	19	16	20	22	77	19.25	120.3
நவம்பர்	22	18	22	24	86	21.50	134.4
டிசம்பர்	22	10	24	25	81	20.25	126.6
					மொத்தம்	192.50	

$$\frac{192.5}{12} = 16.04$$ இது பொதுச்சராசரி

மாதாந்திர மொத்தத்தை வருடங்களின் எண்ணிக்கையால் (இங்கு 4) வகுக்க மாதாந்திர சராசரி கிடைக்கும். மாதாந்திர சராசரியைப் பொதுச்சராசரியால் (இங்கு 16.04) வகுத்து 100ஆல் பெருக்கினால் பருவகால மாறுபாட்டுக் குறியீட்டெண் (index) கிடைக்கும்.

மாதாந்திர புள்ளி விபரங்களுக்குப் பதிலாக காலாண்டு புள்ளி விபரங்கள் இருந்தால், முதலில் காலாண்டு மொத்தம் கண்டுபிடிக்க வேண்டும். பிறகு காலாண்டுச் சராசரி (காலாண்டு மொத்தம் ÷ வருடங்களின் எண்ணிக்கை) கண்டுபிடிக்க வேண்டும். காலாண்டுச் சராசரிகளின் மொத்தத்தை நான்கால் (காலாண்டு என்பதால், ஒவ்வொரு ஆண்டுக்கும் நான்கு புள்ளிகள் இருக்கும் என்பதால்) வகுத்தால் பொதுச்சராசரி கிடைக்கும். பின்னர் ஒவ்வொரு காலாண்டுச் சராசரியையும் பொதுச்சராசரியால் வகுத்து 100ஆல் பெருக்கினால் பருவகால மாறுபாட்டுக் குறியீட்டெண்கள் கிடைக்கும்.

நகரும் சராசரி முறையில் பருவகால மாறுபாட்டுக் குறியீட்டெண்கள் காணல்

முதலில் நகரும் சராசரி கண்டுபிடிக்க வேண்டும். கொடுக்கப்பட்ட மதிப்பை அதற்குரிய நகரும் சராசரியால் வகுத்து அதை 100ஆல் பெருக்கிக் கொள்ள வேண்டும். இதைப் பின்னர் பொதுச்சராசரியால் வகுக்க வேண்டும். வரும் விடையை 100ஆல் பின்னர் பெருக்கினால் கிடைப்பது பருவகால மாறுபாட்டுக் குறியீட்டெண். ஓர் உதாரணம் இங்கு காணலாம். அட்டவணை 75யைப் பார்க்கவும்.

அட்டவணை 75இல் விடப்பட்டுள்ள மதிப்புக்களை எளிதாகக் கண்டுபிடித்துவிடலாம். அந்த மதிப்புக்களையும் சேர்த்து அட்டவணை 76யை உருவாக்கலாம். அதன் கடைசி நிரலில் இருப்பது பருவகால மாறுபாட்டுக் குறியீட்டெண்கள் ஆகும்.

அட்டவணை - 75

ஆண்டு	காலாண்டுப் பகுதி	மதிப்பு	நான்கு காலாண்டு நகரும் மொத்தம்	காலாண்டு நகரும் சராசரி	மையப் படுத்தப் பட்ட நகரும் சராசரி	மதிப்பீடு செய்யப் பெற்ற பருவகால மாறுபாடுகள்
2005	I	30				
	II	40				
			140	35.0		
	III	36			35.5	(36/35 5) × 100 = 101.4
			144	36.0		
	IV	34			37.5	(34/37.5) × 100 = 90.67
			156	39.0		
2006	I	34			40.75	(34/40.75) ×100 = 83.45
			170	42.5		
	II	52			43.75	(52/43.75) ×100 =118.9
			180	45.0		
	III	50			45.75	(50/45.75) ×100 = 109.2
			186	46.5		
	IV	44			47.25	(44/47 25) ×100 =93.13
			192	48		
2007	I	40			48.5	(40/48 5) ×100 =82.49
			196	49		
	II	58			49.5	
			200	50		
	III	54			51 75	
			214	53.5		
	IV	48			55.75	
			232	58		
2008	I	54			59.75	
			246	61.5		
	II	76			63.25	
			260	65		
	III	68				
	IV	62				
2009	I	80				
	II	92				
	III	86				
	IV	82				

அட்டவணை - 76

காலாண்டுப் பகுதி	2005	2006	2007	2008	2009	காலாண்டு மொத்தம்	காலாண்டு சராசரி	பருவகால மாறுபாட்டுக் குறியெண்
I	--	83.45	82.49	90.36	102.9	359.20	89.80	90.05
II	--	118.90	117.1	120.1	111.5	467.60	116.90	117.2
III	101.4	109.2	104.3	99.63	-	414.53	103.63	103.9
IV	90.67	93.13	86	84.35	-	354.23	88.80	88.80
						மொத்தம்	398.88	

மேலே கொடுக்கப்பெற்றுள்ள முறையில் ஒரு காலத்திற்கான மதிப்பினை அக்காலத்திற்கான நகரும் சராசரியால் வகுத்து அதை 100ஆல் பெருக்கிவிட்டால் மதிப்பீடு செய்யப் பெற்ற பருவகால மாறுதல்கள் கிடைக்கின்றன. இதற்குப் பதிலாக, ஒரு காலத்திற்கான மதிப்பிலிருந்து அதன் நகரும் சராசரியைக் கழித்தும் மதிப்பீடு செய்யப்பெற்ற பருவகால மாறுதல்கள் பெறலாம்.

இவ்விரண்டு முறைகளன்றி, இன்னுமொருமுறையும் உள்ளது. அது குறைந்த வர்க்க முறையைப் பயன்படுத்தி போக்கு மதிப்பைக் (Trend value) கண்டுபிடித்து அந்த மதிப்புக்களால் அவற்றிற்குரிய மதிப்புக்களை வகுத்துக் கண்டுபிடிப்பதாகும்.

$$\frac{\text{உண்மையான மதிப்பு}}{\text{போக்கு மதிப்பு}} \quad \frac{\text{(Actual value)}}{\text{(Trend value)}} \times 100$$

இதுபற்றி மேலும் விபரங்களுக்கு, முர்ரே ஆர்.ஸ்பீஜெல் (MURRAY R. SPIEGEL, Schaum's Outline of Theory and Problems of Statistics, McGraw - Hill International Book Company, Singapore, 1961, pp.295-298) எழுதியுள்ள புள்ளியியல் புத்தகத்தைப் பார்க்கலாம்.

பருவகால மாறுபாடுகளகற்றுதல் (Deseasonalisation)

பருவகால மாறுபாட்டுக் குறியீட்டெண்களைப் பெற்ற பின்னர் (விவரிக்கப்பட்டுள்ள முறைகளில் ஏதேனும் ஒன்றின் மூலம்) அவற்றைக் கொண்டு பருவகால மாறுபாடுகளை நீக்க வேண்டும். அதற்குப் பருவகால மாறுபாடுகளகற்றுதல் (deseasonalisation of data) என்று பெயர்.

பருவகால மாறுபாடுகளை அகற்றுவதற்குக் கீழ்க்காணும் செயல்பாடுகளைச் செய்ய வேண்டும். உதாரணமாக, நமக்கு ஒவ்வொரு மாதத்திற்கும் பருவகால மாறுபாட்டுக் குறியீட்டெண்கள் கிடைத்துள்ளதென வைத்துக் கொள்வோம். 2000 முதல் 2009 வரையிலான மாதாந்திரப் புள்ளி விபரங்கள் உள்ளனவென்றும் கொள்வோம். 2000 முதல் 2009 வரையிலான ஜனவரி மாத மதிப்புகளை ஜனவரி மாதத்திற்கான பருவகால மாறுபாட்டுக் குறியீட்டெண்ணால் வகுத்தால் கிடைப்பது பருவகால மாறுபாடுகள் அகற்றப்பட்ட மதிப்புக்களாகும். அதேபோல், பிப்ரவரி மாத மதிப்புக்களை பெப்ரவரி மாத பருவகால மாறுபாட்டுக் குறியீட்டெண்ணால் வகுக்க பிப்ரவரி மாதத்திற்கான பருவகால மாறுபாடுகள் அகற்றப்பட்ட மதிப்புக்கள் கிடைக்கும். இவ்வாறு எல்லா மாதங்களுக்கும் செய்து மொத்தமாகப் பருவகாலத்தினால் ஏற்பட்ட மாறுபாடுகளை அகற்றி விடலாம்.

சுழல் ஏற்ற இறக்கங்களைக் கணித்தல்
(Measurement of Cyclical variations)

இதுவரை நீண்டகாலப்போக்கு (T) மற்றும் பருவகால மாறுபாடுகளைக் (S) கணிப்பதைப் பற்றி விளக்கப்பட்டது. இவைகளைக் கணித்த பின்னர் சுழல் ஏற்ற இறக்கங்களை எளிதில் கணித்துவிடலாம். **பெருக்கல் முறையை** ($Y = T \times C \times S \times I = TCSI$) அடிப்படையாகக் கொண்டால் கீழ்க்கண்டவாறு சுழல் மாறுபாடுகளைக் கணிக்கலாம்.

1. நகரும் சராசரி முறையையோ குறைந்த வர்க்க முறையையோ பயன்படுத்தி நீண்டகாலப் போக்கு (T) மதிப்புக்களை முதலில் கணிக்க வேண்டும்.

2. ஏதேனும் ஒரு முறையைப் (எளிய சராசரி அல்லது நகரும் சராசரி முறை) பயன்படுத்தி பருவகால மாறுபாட்டுக் குறியீட்டெண்களைப் (S) பெற வேண்டும்.

3. கொடுத்துள்ள தொடர்வரிசையில் உள்ள ஒவ்வொரு மதிப்பையும் (Y) அவற்றிற்கெதிரில் உள்ள Tன் மதிப்பு Sன் மதிப்பு ஆகியவற்றின் பெருக்குத் தொகையால் வகுக்க வேண்டும் (TS). அதாவது $\frac{Y}{TS}$. பெருக்கல் முறைப்படி Y = TCSI; $\frac{Y}{TS}$ = CI. இதில் C என்பது சுழல் மாறுபாடுகளையும், I (irregular) என்பது ஒழுங்கற்ற எதிர்பாராத ஏற்ற இறக்கங்களையும் குறிக்கின்றன.

4. (C×I) மதிப்புக்களின் வரிசையிலிருந்து நகரும் சராசரியைக் கணிக்க வேண்டும். நகரும் சராசரி மதிப்புக்களே சுழல் ஏற்ற இறக்கங்களின் மதிப்புக்களாகும். ஏனெனில், நகரும் சராசரியைக் கணிப்பதன் மூலம் ஒழுங்கற்ற ஏற்ற இறக்கங்களாகிய I என்ற பகுதியை (C × I) மதிப்புக்களிலிருந்து நீக்கிவிட முடிகிறது.

கூட்டல் முறையை அடிப்படையாகக் கொண்டால் சுழல் மாறுபாடுகளைக் கீழ்க்கண்டபடி பெறலாம்.

1. முதலில் நீண்டகாலப் போக்கு மதிப்புக்களை (T) கணிக்க வேண்டும்.

2. கூட்டல் முறையை அடிப்படையாகக் கொண்டு நகரும் சராசரி முறையைக் கையாண்டு பருவகால மாறுபாடுகளை (S) கணிக்க வேண்டும்.

3. பின்னர், கொடுத்துள்ள தொடர்வரிசையில் உள்ள ஒவ்வொரு (Y) மதிப்பிலிருந்தும் அதற்கெதிரில் உள்ள T, S ஆகிய இரு மதிப்புக்களையும் கழிக்க வேண்டும். Yலிருந்து T+Sயைக் கழித்து விட்டதால், மீதமிருப்பது C+I ஆகும்.
4. C+I மதிப்புக்களிலிருந்து நகரும் சராசரியைக் கணிப்பதன் மூலம், ஒழுங்கற்ற ஏற்ற இறக்கங்களை C+Iலிருந்து நீக்கிவிட முடியும்.
5. (C×I) அல்லது (C+I) மதிப்புக்கள் மிகவும் ஒழுங்கற்ற முறையில் அமைந்திருந்தால் நகரும் சராசரியின் கால அளவைக் குறைவாகத் தேர்ந்தெடுத்துக் கொள்ள வேண்டும்.

ஒழுங்கற்ற ஏற்ற இறக்கங்களைக் கணித்தல் (Measurement of irregular fluctuations)

கொடுக்கப்பட்டுள்ள தொடர்வரிசை மதிப்புக்களிலிருந்து நீண்டகாலப்போக்கு, பருவகால மாறுபாடுகள், சுழல் ஏற்ற இறக்கங்கள் ஆகிய மூன்றினையும் நீக்கிய பின்னர் எஞ்சி இருக்கும் மதிப்புக்களே ஒழுங்கற்ற ஏற்ற இறக்கங்களின் மதிப்புகளாகும். ஒழுங்கற்ற ஏற்ற இறக்கங்கள் மிகக் குறைந்த அளவில் அடிக்கடியும் மிக அதிக அளவில் எப்போதாவதும் நிகழக்கூடியவையாக இருக்கும்.

முன்கணிப்பு (Forecasting)

ஒரு நீண்ட நேர்கோட்டில் ஏதேனும் ஒரு பகுதியை வெட்டி எடுத்துவிட்டாலும், மீதமுள்ள இரண்டு பகுதிகளை வைத்து இடையில் வெட்டி எடுக்கப்பட்ட பகுதியை வரைந்து விடலாம். இதற்கு இடைச் செருகுதல் (interpolation) என்று பெயர். அதுபோல, கடைசியில் இருக்கும் பகுதியை இன்னும் சில ஆண்டுகளுக்கு நீட்டவும் (extrapolate) முடியும். இடைச்செருகுதலும் (interpolation) நீட்டுதலும் (extrapolation) நேர்கோடுகளுக்கு எளிதாகவும் சரியாகவும் இருக்கும்.

வளைகோடுகளுக்கு இம்முறைகள் மூலம் சரியான புள்ளி விபரங்கள் கண்டுபிடிப்பது சற்று சிரமமாக இருக்கலாம். கிடைத்த புள்ளிகளைக் கொண்டு இனிமேல் என்ன நடக்கலாம் என்பது முன்கணிப்பு (FORECASTING) என அழைக்கப்படுகிறது. அதுபோல் பல காலங்களுக்குப் பின்னால் என்ன நடந்திருக்கும் எனப் பின்கணிப்பும் (BACKCASTING) செய்யலாம். பொதுவாக வரலாற்றுத் துறையினருக்கு பின்கணிப்பு மிக அவசியமாகத் தேவைப்படலாம். இதனாலேயே, வரலாற்றுத்துறை வல்லுநர் (R.A. FOGEL) ஒருவருக்கு அவர் பயன்படுத்திய முறைக்காக (CLIOMETRICS) நோபல் பரிசு கிடைத்தது. முன்கணிப்பு செய்வது அனைவருக்கும் அவசியமாக உள்ளது. மழையின் அளவு, காற்றின் அளவு, பங்குச் சந்தை நிலவரம் போன்றவற்றைப் பற்றி முன்னமே அறிவது பல நன்மைகளைத் தரலாம். சுனாமி (TSUNAMI) வருவதைப் பற்றிய முன்னறிவிப்பு நிறைய இழப்புக்களைத் தவிர்க்கப் பயன்படலாம்.

முன்கணிப்பு பலவகைப்படலாம். எந்த மாற்றங்களையும் செய்யாமல் எதிர்காலத்தில் என்ன நடக்கும் என முன்கணிப்புச் செய்யலாம். ஏதேனும் சிறிய / பெரிய மாற்றங்களைச் செய்து அதனால் அடுத்து என்ன நிகழும் எனவும் முன்கணிப்புச் செய்யலாம். நீண்டகால முன்கணிப்பும் குறுகிய கால முன்கணிப்பும் உள்ளன. முன்கணிப்பின் மேல் தாக்கம் ஏற்படுத்தும காரணிகளைப் பொறுத்தும் முன்கணிப்புகள் வகைப்படுத்தப்படலாம். உதாரணமாக, அரசியல், பொருளாதார, சமூக முன்கணிப்புகளும் உள்ளன. முன்கணிப்பு செய்யும் அளவினைப் பொறுத்து, நுண்மை (Micro) முன்கணிப்பு பெரும் (macro) முன்கணிப்பு என்றும் கூறப்படுகின்றன.

கீழேகொடுக்கப்பட்டுள்ள பொழுதுபோக்குப் பொருட்களுக்கான உற்பத்திச் செலவைப் பயன்படுத்தி (ரூபாய் கோடியில்) 2010ஆம் ஆண்டுக்கான செலவை எப்படி முன்கணிப்பு செய்யலாம் என்று பார்க்கலாம்.

அட்டவணை - 77

	ஜன	பிப்	மார்ச்	ஏப்	மே	ஜூன்	ஜூலை	ஆக	செப்	அக்	நவ	டிச
2002	318	281	278	250	231	216	223	245	269	302	325	347
2003	342	309	299	268	249	236	242	262	288	321	342	364
2004	367	328	320	287	269	251	259	284	309	345	367	394
2005	392	349	342	311	290	273	282	305	328	364	389	417
2006	420	378	370	334	314	296	305	330	356	396	422	452
2007	453	412	398	362	341	322	335	359	392	427	454	483
2008	487	440	429	393	370	347	357	388	415	457	491	516
2009	529	477	463	423	398	380	389	419	448	493	526	560

பொதுவாகக் காலம் சார்ந்த தொடர்பான மாறிகளின் மதிப்புக்கள் பலவகையான (பருவ, சுழல், ஒழுங்கற்ற) மாறுபாடுகளாலும் ஏற்ற இறக்கங்களாலும் பாதிக்கப்படுவதால், நீண்டகாலப் புள்ளி விபரங்களைக் கொண்டு முன்கணிப்புச் செய்வது நல்லது. மாறாக, சில புள்ளிவிபரங்களைக் கொண்டு முன்கணிப்பு செய்யும்போது அது தவறாக முடியலாம். ஏனெனில், நமக்குக் கிடைத்த சில புள்ளிவிபரங்கள் சுழலின் ஏதேனும் ஒரு பகுதியில், நிலையில், கட்டத்தில் இருக்கலாம். அடுத்த சில புள்ளிகளின் போக்கு வேறு கோணத்தில் மாறலாம். எனவே குறைந்தது ஒரு முழுச் சுழல் பற்றிய புள்ளிவிபரங்களைக் கொண்டு அடுத்துவரும் புள்ளிவிபரங்களை முன்கணிப்பு செய்வது நல்லது.

கொடுக்கப்பட்டுள்ள புள்ளிவிபரங்களைக் கொண்டு 2010ஆம் ஆண்டுக்கான உற்பத்திச் செலவை முன்கணிப்புச் செய்ய கீழ்வரும் முறைகளை ஒவ்வொன்றாகச் செய்யலாம். முதலில் 12 மாத மையப்படுத்தப்பட்ட நகரும் சராசரியைக் காணலாம். இதில், எல்லா மாதங்களுக்கும் உள்ள அசையும் சராசரியை இங்கு கொடுப்பதற்கு அதிக இடம் தேவைப்படும் என்பதால் ஒரு பகுதி மட்டும் கொடுக்கப்படுகிறது. ஜூலை மாதம் 2007லிருந்து ஜூன் மாதம் 2009 வரையிலான அசையும் சராசரியை அட்டவணை 78ல் காணலாம்.

அட்டவணை - 78

மாதம்	ஆண்டு	12 மாத மையப்படுத்தப் பட்ட நகரும் சராசரி	மாதம்	ஆண்டு	12 மாத மையப்படுத்தப் பட்ட நகரும் சராசரி
ஜூலை	2007	396.2	ஜூலை	2008	425.9
ஆக	2007	398.8	ஆக	2008	429.2
செப்	2007	401.3	செப்	2008	432.2
அக்	2007	403.9	அக்	2008	434.8
நவ	2007	406.4	நவ	2008	437.2
டிச	2007	408.6	டிச	2008	439.8
ஜன	2008	410.6	ஜன	2009	442.5
பிப்	2008	412.7	பிப்	2009	445.1
மார்ச்	2008	414.9	மார்ச்	2009	447.8
ஏப்	2008	417.1	ஏப்	2009	450.7
மே	2008	419.9	மே	2009	453.6
ஜூன்	2008	422.8	ஜூன்	2009	456.9
	மொத்தம்	4913.2		மொத்தம்	5295.7
	சராசரி	409.4		சராசரி	441.3

மேலே கொடுக்கப்பட்டுள்ள புள்ளி விபரங்களிலிருந்து 2 புள்ளிகள் கிடைக்கின்றன. அவ்விரண்டு புள்ளிகளையும் தொட்டுச் செல்லுமாறு ஒரு நேர்கோடு வரைந்தால் ஜூலை 2009 முதல் டிசம்பர் 2010 வரைக்கான முன்கணிக்கப்பட்ட புள்ளி விபரங்கள் கிடைக்கும்.

இன்னொரு முறையிலும் முன்கணிப்பு செய்யலாம். முதலில் 2007-2008 மற்றும் 2008-2009க்கான இரண்டு சராசரிகளையும் எடுத்துக்கொள்ள வேண்டும். அவற்றிற் கிடையே உள்ள வித்தியாசம் 12 மாதங்களுக்கு உரியதாகையால், அந்த வித்தியாசத்தை 12ஆல் வகுத்து விட்டால் ஒவ்வொரு மாதமும் ஏற்பட்டுள்ள மாற்றம் தெரியும். அந்த மாற்றத்தைக் கொண்டு தேவையான மாதங்களுக்கு முன்கணிப்பு செய்து போக்கு மதிப்பினைப் பெறலாம். இங்கு கிடைத்துள்ள சராசரிகள் 441.3 மற்றும் 409.4. இவற்றிற்கிடையேயான வித்தியாசம் 31.9. இதை 12ஆல்

வகுக்கக் கிடைப்பது 2.66. முதலில் 2009 ஜூன் மாதத்திற்கான அசையும் சராசரியான 456.5 உடன் 2.66ஐக் கூட்டினால் 2009 ஜூலை மாதத்திற்கான போக்கு (Trend) மதிப்பினைக் காணலாம். இவ்வாறாக டிசம்பர் 2010 வரை ஒவ்வொரு மாதத்திற்கும் காணலாம். இவற்றை அட்டவணை 79இல் காணலாம்.

அட்டவணை - 79

	ஜன	பிப்	மார்ச்	ஏப்	மே	ஜூன்	ஜூலை	ஆக	செப்	அக்	நவ	டிச
2009க்கான போக்கு மதிப்பு						456.9	459.6	462.2	464.9	467.5	470.2	472.9
2010க்கான போக்கு மதிப்பு	475 5	478 2	480 8	483.5	486.2	488.8	491.5	494.1	496.8	499.5	502.1	504.8

அட்டவணை 79இல் உள்ள 2010க்கான போக்கு மதிப்புக்களை அந்தந்த மாதத்திற்கான பருவகால குறியீட்டெண்ணுடன் பெருக்கினால் முன்கணிப்பு கிடைக்கும். பருவகாலக் குறியீட்டெண்கள் பெற முதலில், 12 மாத மையப்படுத்தப்பட்ட நகரும் சராசரி காணவேண்டும். அவ்வாறு வகுத்துக் கண்டுபிடிக்கப்பட்ட முடிவுகள் அட்டவணை 80இல் தரப்படுகின்றன. உதாரணமாக, 2002ஆம் ஆண்டு ஜூலை மாதத்திற்கான $81.2 = (223 \div 274.7) \times 100$. இதில் 274.7 என்பது ஜூலை 2002க்கான நகரும் சராசரி.

அட்டவணை 80இல் உள்ள கடைசி நிரையில் இடைநிலை பயன்படுத்தப்பட்டுள்ளது. ஏனெனில், ஒவ்வொரு மாதத்திற்கும் உள்ள மதிப்புகள் அதிக வித்தியாசம் உள்ளவைகளாக இருக்கின்றன. இடைநிலைக்குப் பதிலாக கூட்டுச் சராசரியும் பயன்படுத்தப்படலாம். அட்டவணை 80இல் கடைசி நிரையில் உள்ளவைதான் அந்தந்த மாதத்திற்கான பருவகாலக் குறியீட்டெண்கள். இந்தக் குறியீட்டெண்களைக் கொண்டு, அட்டவணை 79இல் உள்ள மூன்றாவது நிரையில் உள்ள போக்கு மதிப்புகளைப் பெருக்கி

100ஆல் வகுத்து விட்டால், கிடைப்பது அந்தந்த மாதத்திற்கான கணிக்கப்பட்ட (predicted) மதிப்புகள்.

அட்டவணை - 80

ஆண்டு	ஜன	பிப்	மார்ச்	ஏப்	மே	ஜூன்	ஜூலை	ஆக	செப்	அக்	நவ	டிச
2002	-	-	-	-	-	-	81.2	88.5	96.4	107.6	115.2	122.3
2003	119.9	107.7	103.7	92.4	85.4	80.6	82.2	88.4	96.6	107.1	113.5	120.2
2004	120.7	107.3	104.1	92.8	86.4	80.0	82.0	89.3	96.7	107.2	113.4	121.1
2005	119.8	106.1	103.4	93.6	86.8	81.3	83.4	89.6	95.7	105.5	112.2	119.6
2006	119.8	107.2	104.3	93.5	87.2	81.5	83.4	89.5	95.9	106.0	112.3	119.6
2007	119.1	107.6	103.2	93.2	87.2	81.8	84.6	90.0	97.7	105.7	111.7	118.2
2008	118.6	106.6	103.4	94.2	88.1	82.1	83.8	90.4	96.0	105.1	112.3	117.3
2009	119.5	107.2	103.4	93.9	87.7	83.2	-	-	-	-	-	-
இடை நிலை	119.8	107.2	103.4	93.5	87.2	81.5	83.4	89.5	96.4	106.0	112.3	119.6

இவ்வாறாக, கணிக்கப்பட்ட மதிப்புகள் பெறுவதற்கு கொடுக்கப்பட்ட மதிப்புகளின் நகரும் சராசரி, பருவகாலக் குறியீட்டெண், போக்கு மதிப்பு (trend value) ஆகியவை பெற வேண்டும். 2010ஆம் ஆண்டுக்கான முன்கணிப்பு மதிப்புகள் அட்டவணை 81இல் தரப்படுகின்றன.

அட்டவணை - 81

மதிப்புகள்	ஜன	பிப்	மார்ச்	ஏப்	மே	ஜூன்	ஜூலை	ஆக	செப்	அக்	நவ	டிச
2010க்கான போக்கு மதிப்பு	475.5	478.2	480.8	483.5	486.2	488.8	491.5	494.1	496.8	499.5	502.1	504.8
பருவகாலக் குறியீட்டெண் (%)	119.8	107.2	103.4	93.5	87.2	81.5	83.4	89.5	96.4	106.0	112.3	119.6
2010க்கான முன் கணிக்கப்பட்ட மதிப்பு	570	513	497	452	424	398	410	442	479	529	564	604

குறிப்பு : பருவகாலக் குறியீட்டெண் சதவீதத்தில் உள்ளது. எனவே முன்கணிக்கப்பட்ட மதிப்புகளைப் பெற 100ஆல் வகுக்க வேண்டும்.

தொடர்சார்பு

ஒரு சார்பு மாறி பல மாறிகளைச் சார்ந்திருக்கலாம். அவை அதே காலத்தைச் சார்ந்த மாறிகளாகவோ, அதற்கு முந்திய காலத்தைச் சார்ந்த மாறிகளாகவோ இருக்கலாம்.

உதாரணத்திற்கு, ஒரு நிறுவனத்தின் இந்த ஆண்டுக்கான வருமானம் (Y_t) சென்ற ஆண்டு செய்யப்பட்ட விளம்பரத்தைப் (A_{t-1}) பொறுத்து அமையலாம். அதுபோல, இந்த ஆண்டின் வருமானம் (Y_t) சென்ற ஆண்டின் வருமானத்தைப் (Y_{t-1}) பொறுத்தும் அமையலாம். இதில் Y_t என்பது வெளிமாறி (exogenous variable) எனவும் Y_{t-1} உள்மாறி (endogenous variable) எனவும் அழைக்கப்பெறுகின்றன. கால இடைவெளி குறைவாக இருந்தால் Y_tக்கும் Y_{t-1}க்கும் நெருக்கமான உறவு இருக்கலாம். இவை இரண்டிற்கும் இடையே உள்ள உறவை தொடர்உறவு (serial correlation) என்கிறார்கள்.

$$r_1 = \frac{Cov\,(Y_t\, Y_{t-1})}{\sqrt{Var\,(Y_t)\; var\,(Y_{t-1})}}$$

Y_t க்கும் Y_{t-1} க்கும் இடையே ஒரு கால அலகு இருப்பதால், இது ஒரு படி (First order) தொடர் உறவு என அழைக்கப் படுகிறது. இதனையே சிலர் தற்சார்பு (antocorrelation) எனவும், பின்தொடரும் உறவு (lag correlation) எனவும் அழைக்கின்றார்கள். ஆனாலும் சிலர் இவற்றிற்கிடையே சிறிதளவு வித்தியாசம் இருப்பதாகவும் கூறுகிறார்கள். ஒரு முழுமையில் உள்ள புள்ளிகளுக்கு இடையே உள்ள தொடர் உறவைத் தற்சார்பு (autocorrelation) எனவும், ஒரு மாதிரியில் உள்ள புள்ளிகளுக்கு இடையே உள்ள தொடர் உறவைத் தொடர் உறவு (serial correlation) எனவும், ஒரே மாறியின் மதிப்பு வெவ்வேறு காலத்தில் உள்ள மதிப்புக்களுடன் உடன்தொடர்பு கொண்டிருந்தால் அது பின்தொடரும் உறவு (lag correlation) எனவும் அழைக்கப்படுகின்றன.

காலம் சார்ந்த மதிப்புகள் மாறிக் கொண்டிருக்கும் தன்மை கொண்டவை. அவை மாறுகின்ற தன்மையை வைத்து அசையும் தொடர்வரிசை (nonstationary) எனவும் அசையாத் தொடர்வரிசை (stationary) எனவும் அழைக்கப் பெறுகின்றன. அசையாத தொடர்வரிசையும் ஏற்ற இறக்கங்களைக் கொண்டிருக்கலாம். அந்த ஏற்ற இறக்கங்கள் தற்செயலாக நடந்தனவா அல்லது சுழல் இயக்கத்தால் (cyclical) அல்லது

அலைவியக்கத்தால் (oscillatory) நடந்தனவா என்பதைக் கவனிப்பது அவசியம். ஏனெனில், அதைப் பொறுத்து பிழைகளுக்கு (errors) இடையே உள்ள உறவும் அமைகின்றது; அதைப் பொறுத்து பண்பலகுகளின் (parameters) புள்ளியியல் முக்கியத்துவதைச் சோதனை செய்வதும் அமைகிறது. பல பொருளியல் நிகழ்வுகள் அலைவியக்கத்தால் பாதிக்கப் பெறுகின்றன. உதாரணமாக, சந்தைப் பொருளியலில் பொருளாதாரக் குழப்பங்கள் (economic crisis) என்பது அலைவியக்கத்தால் உருவாகின்றன எனலாம். 2008-2009ஆம் ஆண்டு பொருளாதாரக் குழப்பம் கூட அத்தகையதே.

காலம்சார் தொடர் வரிசையின் போக்கினை அறுத்த (detrending) பின்னர், அது அசைவற்ற (stationary) நிலையை அடைந்துள்ளதா எனப் பார்த்து அதனுடைய அமைப்பினை (form) முடிவு செய்ய வேண்டும். உதாரணமாக, X_t, X_{t-1}, X_{t-2}, ... போன்ற அசைவற்ற காலம்சார் தொடர்வரிசை இருந்தால், அதை

$$X_t = aX_{t-1} + b + e_t$$ எனலாம்.

இதில் aயும் bயும் மாறிலிகள். e_t தற்செயலாக சந்தர்ப்பவசத்தால் வந்த இடையூறு (random disturbance) ஆகும். இதில், X ஓர் ஆண்டு பின்தங்கிய மாறியான X_{t-1} உடன் தொடர்பு கொண்டுள்ளதால், இதனை, ஒருபடி (first order) தொடர் தற்சார்பு (auto regressive) எனலாம். இதனை, ஒருபடி வித்தியாசச் சமன்பாடு (first order difference equation) எனவும் அழைக்கலாம். இதுபோல இரண்டுபடி (second order), மூன்றுபடி .... சமன்பாடுகளையும் ஆய்வுகளில் காணலாம். இப்படிப்பட்ட பலவகையான போக்குகளையும் உறவுகளையும் கையாளத் தெரிந்தால், நம்பகத்தன்மைமிக்க பயனுள்ள ஆய்வுகளைச் செய்யமுடியும்.

அசைவின்மைக்கான சோதனைகள் (Tests for stationarity)

மாறிகளைப் பலவகைகளாக முன்னர் பிரித்துள்ளோம்.

மாறும் தன்மையினை மையமாக வைத்து, மாறிகளை இரண்டு வகைகளாகப் பிரிக்கலாம். அவை, அசையும் (non-stationary) மாறிகள் மற்றும் அசையா மாறிகள் (stationary) ஆகும். மாறிகளின் போக்கும் குணங்களும் அவற்றின் அசையும் தன்மையால் பாதிக்கப்படுகின்றன. ஒரு மாறியின் மதிப்புக்கள் ஒரு நிலைநிறுத்தப்பட்ட சராசரியின் (fixed mean) மதிப்பைச் சுற்றி மேலும் கீழும் மட்டும் அசைந்து கொண்டிருந்தால் அந்த வகை மாறி அசையா மாறி என அழைக்கப்படுகிறது. மாறாக, ஒரு குறிப்பிட்ட சராசரியைச் சுற்றி அமையாமல், காலப்போக்கில் ஒரு மாறியின் மதிப்புக்கள் ஏதேனும் ஒரு திசையில் (மேலோ, கீழோ) அசைந்து சென்று கொண்டிருந்தால் அது அசையும் மாறி (non-stationary) என அழைக்கப்பெறுகிறது. இந்நிலை பரவலாக காலம்சார் தொடர் மாறிகளில் காணப்படுகின்றது. இப்படிப்பட்ட அசையும் மாறிகளுக்கிடையேயான உறவுகள் (correlation, regression) புள்ளியியல் முக்கியத்துவம் பெற்றுள்ளனவா என்று சோதிப்பது சிரமம். ஏனெனில் அப்படிப்பட்ட பண்பலகுகள் (parameters) எவ்விதக் கோட்பாட்டுப் பரவலுக்குள்ளும் வராமல் இருக்கும். இதற்கு எடுத்துக்காட்டாக, பொய்யான அல்லது போலியான உடன்தொடர்பினைக் (spurious regression) கூறலாம். எனவே மாறிகளுக்கிடையே உடன் தொடர்பு உள்ளதா? அந்த உடன் தொடர்பு புள்ளியியல் முக்கியத்துவம் பெற்றதா? என்று காண்பதற்கும் முன்பாக, எடுத்துக்கொண்டுள்ள மாறிகள் அசையும் தன்மை உடையனவா, அல்லது அசையாத் தன்மையைப் பெற்றனவா என்று கணிப்பது அவசியமாகிறது. அசையாத்தன்மை கொண்ட மாறிகளையே ஆய்வுக்கு உட்படுத்துவது சிறந்தது. அசையும் தன்மை உடைய மாறிகளாக இருந்தால் அந்த மாறிகளின் அசையும் தன்மையை அப்புறப்படுத்திவிட்டு அந்த மாறிகளை ஆய்வுக்குப் பயன்படுத்தலாம். அசையும் தன்மை உள்ளதா இல்லையா எனக் காண சில சோதனைகள் உள்ளன. அனுபவம் மிகுந்தவர்களால் புள்ளி விபரங்களின் போக்கினை வைத்தே

அவை அசையும் தன்மை உள்ளனவா இல்லையா எனச் சொல்ல முடியும். அல்லது ஒட்டுறவு வரைபடத்தின் (correlogram) மூலம் சொல்ல முடியும். மிகச் சரியாகக் கணிப்பதற்கு தொடர்போக்கு (auto-regressive) முறைகளைப் பயன்படுத்தும் சோதனைகள் உள்ளன. அவற்றில் சில டிக்கி-ஃபுல்லர் (DICKEY-FULLER) சோதனை, நீட்டிக்கப்பட்ட டிக்கி-ஃபுல்லர் (Augmented Dickey-Fuller) சோதனை, ஃபிலிப்ஸ்-பெர்ரன் (PHILLIPS - PERRON) சோதனை, கிட்டாவஸ்கி - ஃபிலிப்ஸ் - ஸ்மிட் - சின் (KITAWOSKI - PHILLIPS - SCHMIDT - SHIN : KPSS) சோதனை போன்றவை ஆகும். இவற்றைப் பற்றி அதிகமாகத் தெரிந்து கொள்ள இணைய தளத்தைப் பயன்படுத்தலாம். மேலும், Bowerman.B.L. மற்றும் O'Connell R.T. (1979) எழுதியுள்ள (Time Series and Forecasting, Duxbury Press, North Scituate, Massachusetts) புத்தகத்தைப் படிக்கலாம்.

இங்கு கூறப்பட்டுள்ள சோதனைகளில், நீட்டிக்கப்பட்ட டிக்கி-ஃபுல்லர் சோதனை (ADF Test) மிகப் பரவலாகப் பயன்படுத்தப் பெறுகிறது. இந்த சோதனை,

$$\Delta Y_t = b_0+\beta Y_{t-1}+\theta_1\Delta Y_{t-1}+\theta_2\Delta Y_{t-2}+\theta_3\Delta Y_{t-3}+\theta_4\Delta Y_{t-4}+e_t \quad \dots (1)$$

என்ற சமன்பாட்டினை (equation) போக்கு (trend) இல்லாத போதும்,

$$\Delta Y_t = b_0+\beta Y_{t-1}+\theta_1 Y_{t-1}+\theta_2 Y_{t-2}+\theta_3 Y_{t-3}+\theta_4\Delta Y_{t-4}+b_2 t + e_t \quad \dots (2)$$

என்னும் சமன்பாட்டினை போக்கு உள்ளபோதும் பயன்படுத்துகிறது.

இதில் Y என்பது எடுத்துக் கொள்ளப்பட்ட மாறி

$$\Delta Y_t = Y_t - Y_{t-1},$$

t = காலம்

b_0 = மாறிலி (constant)

β, θ, b_2 = போக்குக்கோட்டின் பண்பளவைகள் அல்லது கெழுக்கள்

$\beta = b_2 - 1$

e = பிழை

கொடுக்கப்பட்டுள்ள தொடரில் அசைவு (non-stationarity) அல்லது அசைவின்மை (stationarity) உள்ளதா என்று கணிக்க 't' சோதனையைப் பயன்படுத்தலாம். அதற்கான எடுகோள்கள்:

இல்லெனும் எடுகோள் (H_0) : $\beta = 0$ (அசைவின்மை)

மாற்று எடுகோள் (H_1) : $\beta < 0$ (அசையும் தன்மை)

கணக்கிடப்பட்ட 't'ன் மதிப்புக்கள் எதிரிடையாகவும் (negative) பெரியதாகவும் இருப்பின், இல்லெனும் எடுகோள் மறுக்கப்படும்.

அட்டவணை - 82

ADF சோதனையின் பட்டியல் மதிப்புக்கள்

முக்கியத்துவ அளவு	1%	5%	10%
ADF போக்கு இல்லை (No Trend)	-3.481	-2.884	-2.574
ADF போக்கு உண்டு (Trend)	-4.011	-3.439	-3.139

மேலே கொடுக்கப்பட்டுள்ள அட்டவணை 82இல் இரண்டாவது நிரையில் உள்ள மதிப்புக்களைப் போக்கு (trend) இல்லாதபோதும் (equation 1) மூன்றாவது நிரையில் உள்ள மதிப்புக்களைப் போக்கு உள்ளபோதும் (equation 2) பயன்படுத்த வேண்டும். இந்த ADF சோதனைக்கு கிரெட்ல் (Gretle) என்னும் புள்ளியியல் சிப்பம் (package) பயன்படுத்தப்படலாம்.

ஃ

11. குறியீட்டெண்கள்
(INDEX NUMBERS)

காலத்திற்குக் காலம் விலைகள் மாறலாம்; உற்பத்தி மாறலாம்; நுகர்வு மாறலாம். இந்த மாற்றங்கள் எவ்வளவு என சதவீதத்தில் சொல்லப்படுவது குறியீட்டெண்கள் ஆகும். உதாரணமாக 2008இல் 100ஆக இருந்த ஒரு பொருளின் விலை 2009இல் ரூ.110ஆக கூடினால், அதை 10 சதவீத விலை உயர்வு எனச் சொல்லலாம். அல்லது 100 ஆக இருந்த விலைக் குறியீட்டெண் 110 ஆக ஆகிவிட்டது எனலாம். இந்த 110 = $\frac{110}{100} \times 100$ ஆகும். 2008இல் ரூ.50ஆக இருந்த பொருளின் விலை 2009இல் ரூ.60ஆக மாறினால், 20 சதவீத விலை உயர்வு எனலாம். அதாவது,

$$\frac{60 - 50}{50} \times 100 = 20\%$$ ஆகும்.

இதைக் குறியீட்டெண்ணாகச் சொன்னால் 100 என்று 2008க்கும், 2009க்கு $\left(\frac{60}{50} \times 100\right) = 120$ என்றும் சொல்லலாம். இவற்றில் 100யையும் 120யையும் சதவீத விலைச் சார்பிகள் (percentage price relatives) என்றும் சொல்கிறார்கள். இதில், 60 என்பது நடப்பு ஆண்டு (current year) விலை என்றும் 50 என்பது அடிப்படை ஆண்டு (base year) விலை என்றும் அழைக்கப்படுகின்றன. இக்கருத்துக்களைப் பயன்படுத்தி,

$$\text{விலைக்குறியீட்டெண்} = \frac{\text{நடப்பு ஆண்டு விலை}}{\text{அடிப்படை ஆண்டு விலை}} \times 100$$

$$P_{01} = (P_1 / P_0) \times 100$$

இதேபோல, அளவுக் குறியீட்டெண்கள் (quantity index numbers) மற்றும் மதிப்புக் குறியீட்டெண்கள் (value index numbers) ஆகியவற்றையும் கண்டுபிடிக்கலாம்.

இதுவரை ஒரு காலத்தில் இருந்த ஒரு பொருளின் விலையையோ, ஒரு பொருளின் அளவையோ, ஒரு பொருளின் மதிப்பையோ மற்றொரு காலத்தில் அதன் விலையோடோ, அளவோடோ அல்லது மதிப்போடோ ஒப்பிட்டுப் பார்த்து விலை, அளவு மற்றும் மதிப்பு குறியீட்டெண்கள் கண்டுபிடிப்பது பற்றி விளக்கப்பட்டது. ஒரே சமயத்தில் பல பொருட்களின் விலைகளை அல்லது அளவுகளைக் (quantities) கொண்டு குறியீட்டெண்கள் காண்பது எப்படி என்று இனிக் காணலாம்.

எளிய கூட்டுத் தொகை முறை
(Simple Aggregative Index Numbers)

$$P_{01} = \frac{\Sigma P_1}{\Sigma P_0} \times 100$$

உதாரணமாக, அட்டவணை 83க்கு எளிய கூட்டுத்தொகை முறை மூலம் குறியீட்டெண்ணைக் கணிக்கலாம்.

அட்டவணை - 83

பொருளும் அலகும்	2000ல் விலை (ரூ)	2010ல் விலை (ரூ)
அரிசி (1 கிலோ)	20	25
புகையிலைப் பொருள் (1 கிலோ)	200	400
மதுபானம் (1 லிட்டர்)	150	300
மொத்தம்	370	725

$$\boxed{P_{01}} = \frac{725}{370} \times 100 = \boxed{195.95}$$

இது 2010ஆம் ஆண்டுக்கான குறியீட்டெண்ணாகும்.

சார்பிகளின் எளிய சராசரி முறைக் குறியீட்டெண்
(Simple Average of Relatives)

$$P_{01} = \frac{\Sigma \frac{P_1}{P_0} \times 100}{n}$$

அட்டவணை 83இல் கொடுக்கப்பட்டுள்ள புள்ளி விபரங்களுக்கு சார்பிகளின் எளிய சராசரிமுறையைப் பயன்படுத்தி குறியீட்டெண்ணைக் காணலாம்.

அரிசிக்கான சதவீத விலைச்சார்பி $= \frac{25}{20} \times 100 = 125$

புகையிலைப் பொருளுக்கான

சதவீத விலைச்சார்பி $= \frac{400}{200} \times 100 = 200$

மதுபானத்திற்கான சதவீத விலைச்சார்பி $= \frac{300}{150} \times 100 = 200$

$$\Sigma \frac{P_1}{P_0} \times 100 = 125 + 200 + 200 = 525$$

$$n = 3$$

$$\boxed{P_{01}} = \frac{525}{3} = \boxed{175}$$

சார்பிகளின் பெருக்குச் சராசரி முறைக் குறியீட்டெண்

$$P_{01} = \text{Antilog}\left\{\frac{\Sigma \log \frac{P_1}{P_0} \times 100}{n}\right\}$$

அட்டவணை 83இல் கொடுக்கப்பட்டுள்ள புள்ளி விபரங்களைப் பயன்படுத்தி அட்டவணை 84 பெறப்படுகிறது.

அட்டவணை - 84

பொருள்	P_0	P_1	$\frac{P_1}{P_0} \times 100$	$\log \frac{P_1}{P_0} \times 100$
அரிசி (கிலோ ஒன்றுக்கு)	20	25	125	2.0969
புகையிலைப் பொருள் (கிலோ ஒன்றுக்கு)	200	400	200	2.3010
மதுபானம் (லிட்டர் ஒன்றுக்கு)	150	300	200	2.3010
மொத்தம்				**6.6989**

$$\boxed{P_{01}} = \text{Anti log } (2.23)$$

$$= \boxed{170.9}$$

மூன்று முறைகளுக்கும் பயன்படுத்தப்பட்ட புள்ளி விபரங்கள் மாறவில்லையெனினும் அந்த மூன்று முறைகளால் கண்டுபிடிக்கப்பட்டுள்ள குறியீட்டெண்கள் வித்திய ாக இருப்பதையும், சார்பிகளின் பெருக்குச் சராசரி முறையைப் பயன்படுத்தி, கணித்த குறியீட்டெண் குறைவாக இருப்பதையும் காணலாம்.

விலைக் குறியீட்டெண்களைக் கணித்ததைப் போலவே அளவுக் குறியீட்டெண்களையும் கணிக்கலாம்.

எடைகளின் முக்கியத்துவம்

பொருள்களைப் பல்வேறு பிரிவுகளாகப் பிரிக்கலாம். அனைத்துப் பொருள்களின் விலை மாற்றங்களும் மக்களை ஒரே மாதிரி சமமாகப் பாதிப்பதில்லை. அதுபோல, மக்களையும், அவர்களின் வருமானம், பழக்க வழக்கங்களை வைத்து பல வகைகளாகப் பிரிக்கலாம். ஒவ்வொரு பொருளின் விலை மாற்றமும் வெவ்வேறு வகுப்பினரையும் வெவ்வேறு

விதமாகப் பாதிக்கிறது. எனவே, எல்லாப் பொருள்களுக்கும் சமமான முக்கியத்துவம் கொடுப்பது பொருத்தமில்லாத செயலாகலாம். இச்சூழலில் ஒவ்வொரு பொருளுக்கும் வெவ்வேறு விதமாக முக்கியத்துவம் கொடுப்பது நல்லது.

ஏழைகளின் மொத்த வரவு செலவில் உணவுப் பொருள்களுக்கு அதிக விகிதம் செல்லலாம். செல்வந்தர்களைப் பார்த்து வாழ்க்கைமுறையை மாற்றிக் கொள்ள முனைபவர்கள் போதைப் பொருள்களில் அவர்களின் மிகுதியான வருமானத்தைச் செலவு செய்யலாம். நடுத்தர வர்க்கத்தினர் மருந்து மற்றும் பொழுதுபோக்குப் பொருள்களுக்காகத் தங்கள் வருமானத்தின் கணிசமான தொகையைச் செலவிடலாம். இப்படிப்பட்ட சூழலில், விலை எடைகள் அளவு எடைகள் (quantity weights) மற்றும் மதிப்பு எடைகள் (value weights) ஆகியவை எடைகளாகப் பயன்படுத்தப்படுகின்றன.

எடைகளாகத் தேர்ந்தெடுக்கப்படும் அளவுகள், அடிப்படை ஆண்டைச் சேர்ந்தவைகளாகவோ, நடப்பு ஆண்டைச் சார்ந்தவைகளாகவோ அல்லது அவற்றின் கூட்டு மொத்தமாகவோ இருக்கலாம். வெவ்வேறு வகையான எடைகளுக்குப் பெறும் குறியீட்டெண்கள் வெவ்வேறாக இருக்கலாம். பரவலாகப் பயன்படுத்தப்பட்டு வரும் சில எடையிட்ட குறியீட்டெண்களை இங்கு காணலாம்.

லாஸ்பியரின் குறியீட்டெண் (LASPEYRE'S INDEX NUMBER)

இம்முறை அடிப்படை ஆண்டின் அளவுகளை (Quantities) எடைகளாகப் பயன்படுத்துகிறது.

$$P_{01} = \frac{\Sigma P_1 Q_0}{\Sigma P_0 Q_0} \times 100$$

$$Q_{01} = \frac{\Sigma Q_1 P_0}{\Sigma Q_0 P_0} \times 100$$

பாஷேயின் குறியீட்டெண்
(PAASCHE'S INDEX NUMBER)

இது, விலைக் குறியீட்டெண்ணுக்கு நடப்பு ஆண்டின் அளவுகளை எடையாகவும், அளவுகளின் குறியீட்டெண்ணுக்கு நடப்பு ஆண்டின் விலைகளை எடையாகவும் கொண்டு கணிக்கப்படுவதாகும்.

$$P_{01} = \frac{\Sigma P_1 Q_1}{\Sigma P_0 Q_1} \times 100$$

$$Q_{01} = \frac{\Sigma Q_1 P_1}{\Sigma Q_0 P_1} \times 100$$

ஃபிஷரின் குறியீட்டெண்
(FISHER'S INDEX NUMBER)

ஃபிஷரின் குறியீட்டெண் லாஸ்பியர், பாஷே ஆகியோரின் குறியீட்டெண்களின் பெருக்குச் சராசரியாகும்.

$$P_{01} = \sqrt{\frac{\Sigma P_1 Q_0}{\Sigma P_0 Q_0} \times \frac{\Sigma P_1 Q_1}{\Sigma P_0 Q_1}} \times 100$$

$$Q_{01} = \sqrt{\frac{\Sigma Q_1 P_0}{\Sigma Q_0 P_0} \times \frac{\Sigma Q_1 P_1}{\Sigma Q_0 P_1}} \times 100$$

ஃபிஷரின் குறியீட்டெண் கீழ்க்காணும் காரணங்களினால் விழுமிய குறியீட்டெண் (ideal index number) என அழைக்கப்படுகிறது.

1. அடிப்படை ஆண்டு மற்றும் நடப்பு ஆண்டு ஆகிய இரு ஆண்டுகளின் விலைகளும் அளவுகளும் ஃபிஷரின் குறியீட்டெண்ணில் கணக்கில் கொள்ளப்படுகின்றன.
2. இக்குறியீட்டெண் எவ்வகைப் பிழையும் அற்றது.
3. சிறந்த குறியீட்டெண்ணுக்குரிய சோதனைகளாகிய பொருள், காலம், பகுதி திருப்புச் சோதனைகள் ஆகிய மூன்றினையும் திருப்தி செய்கிறது.

அட்டவணை 85இல் கொடுக்கப்பட்டுள்ள புள்ளிவிபரங்களைக் கொண்டு 2010ஆம் ஆண்டுக்குரிய ஃபிஷரின் விழுமிய விலைக் குறியீட்டெண்ணைக் கணக்கிடலாம். இதற்கு, லாஸ்பியரின் விலைக் குறியீட்டெண்ணும் பாஷேயின் விலைக்குறியீட்டெண்ணும் தேவைப்படுகின்றன. எனவே, முதலில் அவற்றைக் கணிக்கலாம்.

அட்டவணை - 85

பொருள்	அளவு		விலை		Q_0P_0	Q_0P_1	Q_1P_1	Q_1P_0
	Q_0 2000	Q_1 2010	P_0 2000	P_1 2000				
அரிசி	25	30	15	13	375	325	390	450
புகையிலை	10	12	12	20	120	200	240	144
பொழுதுபோக்கு	2	3	60	100	120	200	300	180
மதுபானம்	1	2	20	30	20	30	60	40
	மொத்தம்				635	755	990	814

லாஸ்பியரின் விலைக்குறியீட்டெண் $= \frac{755}{635} \times 100 = 118.89$

பாஷேயின் விலைக்குறியீட்டெண் $= \frac{990}{814} \times 100 = 121.62$

ஃபிஷரின் குறியீட்டெண் $= \sqrt{\text{லா} \times \text{பா}} = \sqrt{118.89 \times 121.62}$

$= 120.25$

மார்ஷல் - எட்ஜ்வொர்த்தின் குறியீட்டெண்
(MARSHALL - EDGEWORTH INDEX NUMBER)

$$P_{01} = \frac{\Sigma P_1(Q_0 + Q_1)}{\Sigma P_0(Q_0 + Q_1)} \times 100$$

$$Q_{01} = \frac{\Sigma Q_1(P_0 + P_1)}{\Sigma Q_0(P_0 + P_1)} \times 100$$

மதிப்புக் குறியீட்டெண் (Value index number)

$$V_{01} = \frac{\Sigma P_1 Q_1}{\Sigma P_0 Q_0} \times 100$$

இது மதிப்பு விகிதம் (value ratio) என்றும் அழைக்கப்படுகிறது.

பிழைகள்

இங்கே நேர்மறை (upward bias) மற்றும் எதிர்மறைப் (downward bias) பிழைகள் என இரண்டு பிழைகள் உள்ளன. விலைக் குறியீட்டெண்ணை அளவுக்குறியீட்டெண்ணால் பெருக்கிவரும் பெருக்கற்பலன் மதிப்புக் குறியீட்டெண்ணாக இருந்தால் அக்குறியீட்டெண்ணை பிழையற்றதாகக் (நேர்மறை, எதிர்மறை ஆகிய இரு பிழைகளுமே இல்லாததாக) கொள்கிறோம்.

உறுதிநிலைக் குறியீட்டெண்கள்
(Fixed base index numbers)

ஒரே ஆண்டை, மீதியுள்ள எல்லா ஆண்டுகளுக்கும் அடிப்படை ஆண்டாகக் கொண்டு கணிக்கப் பெறும் குறியீட்டெண்கள் உறுதிநிலைக் குறியீட்டெண்கள் (fixed base index numbers) என அழைக்கப்படுகின்றன. இவ்வகைக் குறியீட்டெண்களை ஒப்பிடுவது, ஒவ்வொரு ஆண்டிற்கும் வேறுபட்ட ஆண்டை அடிப்படையாகக் கொண்டு கணிக்கப்பெறும் குறியீட்டெண்களை ஒப்பிடுவதைக் காட்டிலும் எளிது.

சங்கிலிக் குறியீட்டெண்கள் (Chain base index numbers)

ஒவ்வொரு ஆண்டிற்கும் அதற்கு முந்தியுள்ள ஆண்டை அடிப்படை ஆண்டாகக் கொண்டு கணிக்கப்பெறும் குறியீட்டெண்களைச் 'சங்கிலிக் குறியீட்டெண்கள்' (Chain base index numbers) என்கிறோம். உதாரணமாக, P_{12} என்பது முதலாவது ஆண்டை அடிப்படையாகக் கொண்டு

இரண்டாவது ஆண்டுக்குக் கணிக்கப் பெற்ற குறியீட்டெண். P_{23} என்றால் இரண்டாவது ஆண்டை அடிப்படை ஆண்டாகக் கொண்டு மூன்றாவது ஆண்டுக்குக் கணிக்கப்பெற்ற குறியீட்டெண்.

உறுதிநிலைக் குறியீட்டெண் தரப்பெற்றிருந்தால், சங்கிலிக் குறியீட்டெண்ணைப் பெறலாம். உதாரணமாக, அட்டவணை 86ஐயக் காணலாம்.

அட்டவணை - 86

தொ. எண்	ஆண்டு	விலை	உறுதிநிலைக் குறியீட்டெண்	சங்கிலிக் குறியீட்டெண்	உறுதிநிலைக் குறியீட்டெண் ணிலிருந்து சங்கிலிக் குறியீட்டெண்
1	2006	50	$(50 \div 50) \times 100 = 100$	100.00	
2	2007	55	$(55 \div 50) \times 100 = 110$	110.00	$(110/100) \times 100 = 110.00$
3	2008	60	$(60 \div 50) \times 100 = 120$	109.09	$(120/110) \times 100 = 109.09$
4	2008	65	$(65 \div 50) \times 100 = 130$	108.33	$(130/120) \times 100 = 108.33$
5	2010	80	$(80 \div 50) \times 100 = 160$	123.08	$(160/130) \times 100 = 123.08$

அட்டவணை 86இல் ஆறாவது நிரலைப் பார்த்தால் உறுதிநிலைக் குறியீட்டெண்ணுக்கும் சங்கிலிக் குறியீட்டெண்ணுக்கும் உள்ள உறவு தெரியவரும். உறுதிநிலைக் குறியீட்டெண்ணிலிருந்து சங்கிலிக் குறியீட்டெண்கள் பெறுவதற்குக் கீழ்க்காணும் விபரங்களைக் காணலாம்.

2009ஐய அடிப்படையாகக் கொண்டு 2010க்குச் சங்கிலிக் குறியீட்டெண் ,

$$P_{2009,\,2010} = \frac{P_{2006,2010}}{P_{2006,2009}} \times 100 = \frac{160}{130} \times 100 = 123.08$$

இதுபோல மற்ற வருடங்களுக்கும் கீழே உள்ளவாறு காணலாம்.

$$P_{2008,\ 2009} = \frac{P_{2006,2009}}{P_{2006,2008}} \times 100 = \frac{130}{120} \times 100 = 108.33$$

$$P_{2007,\ 2008} = \frac{P_{2006,2008}}{P_{2006,2007}} \times 100 = \frac{120}{110} \times 100 = 109.09$$

$$P_{2006,\ 2007} = \frac{P_{2006,2007}}{P_{2006,2006}} \times 100 = \frac{110}{100} \times 100 = 110.00$$

அவற்றிலிருந்து பொதுவான உண்மைகளைப் பெறலாம்.

$$P_{01} = \frac{P_{01}}{P_{00}} \times 100$$

$$P_{12} = \frac{P_{02}}{P_{01}} \times 100$$

$$P_{23} = \frac{P_{03}}{P_{02}} \times 100$$

$$P_{34} = \frac{P_{04}}{P_{03}} \times 100$$

அட்டவணை - 87

ஆண்டு	விலை	உறுதிநிலைக் குறியீட்டெண்	சங்கிலிக் குறியீட்டெண்	சங்கிலிக் குறியீட்டெண்ணிலிருந்து உறுதிநிலைக் குறியீட்டெண்
1	50	100	100.00	
2	55	110	110.00	$110 \times 1.00 = 110$
3	60	120	109.09	$109.09 \times 1.1 \approx 120$
4	65	130	108.33	$108.3 \times 1.09 \times 1.10 \approx 130$
5	80	160	123.08	$123.08 \times 1.08 \times 1.09 \times 1.10 \approx 160$

அட்டவணை 87, அட்டவணை 86லிருந்து பெறப் பெற்றது. அட்டவணை 87லிருந்து சங்கிலிக் குறியீட்டெண்ணைப் பயன்படுத்தி எவ்வாறு உறுதிநிலைக் குறியீட்டெண்ணைப் பெறலாம் என்று தெரியலாம்.

$$P_{15} = P_{45} \times \frac{P_{34}}{100} \times \frac{P_{23}}{100} \times \frac{P_{12}}{100}$$

$$P_{14} = P_{34} \times \frac{P_{23}}{100} \times \frac{P_{12}}{100}$$

$$P_{13} = P_{23} \times \frac{P_{12}}{100}$$

$$P_{12} = P_{12} \times P_{11}$$

வாழ்க்கைச் செலவு குறியீட்டெண்
(Construction of Cost of Living Index Number)

மக்கள் செலவு செய்யும்விதம் அவர்களின் வருமானம் மற்றும் வாழும் இடத்தைப் பொறுத்து அமைகிறது. எனவே, அவர்களின் வாழ்க்கைச் செலவை வெவ்வேறு பொருட்களின் விலைகள் வெவ்வேறு விதமாகப் பாதிக்கின்றன. எனவே, ஒவ்வொரு குழு மக்களுக்கும் வெவ்வேறு குறியீட்டெண்கள் தேவைப்படுகின்றன.

வாழ்க்கைச் செலவுக் குறியீட்டெண்ணை அமைக்கும் முறை

இதற்கு நான்கு முக்கிய படிகள் உள்ளன.

1. கணக்கில் கொள்ள வேண்டிய மக்கள் பிரிவு பற்றி முடிவு செய்தல்.
2. பொருத்தமான பொருட்களைத் தேர்ந்தெடுத்தல்
3. சில்லறை விலைகளைப் பெறுதல்
4. எடைகள் பற்றித் தீர்மானித்தல்

பின்னர், லாஸ்பியரின் சூத்திரத்தைப் பயன்படுத்தி விலைக் குறியீட்டெண்ணைக் கணிக்கலாம்.

இதில் பொருத்தமான அடிப்படை ஆண்டைத் தேர்ந்தெடுப்பதற்கு மிகுந்த முக்கியத்துவம் அளிக்க வேண்டும்; ஏனெனில், அது குறியீட்டெண்ணைப் பாதிக்கும். அடிப்படை ஆண்டு எந்தவித ஏற்ற இறக்கங்களாலும் பாதிக்கப்பட்டு இராத சாதாரண (Normal) ஆண்டாகவும், பொருள்களின் குணங்கள் (Characteristics or attritbutes) மாறியிருக்காத அளவுக்கு அருகில் உள்ள ஆண்டாகவும் இருத்தல் அவசியம். ஒரு தொழிலைச் செய்து ஒரே அளவு வருமானம் பெறும் ஓரிடத்தில் உள்ள விவசாயக் கூலியின் நுகர்வுப் பொருள்களின் விகிதாச்சாரம் காலப்போக்கில் மாறலாம். உதாரணமாக, 20 ஆண்டுகளுக்கு முன் ஒரு விவசாயக் கூலி தண்ணீருக்காகவோ, கைபேசிக் (Cell phone)காகவோ ஏதும் செலவு செய்திருக்கமாட்டார். ஆனால் இன்று அவர் இந்த இரண்டிலும் பணம் அதிகம் செலவு செய்யலாம். அது போல, இன்று கல்லூரியில் படிக்கும் மாணவர்கள் பலர் பாடப்புத்தகங்கள் வாங்காமலேயே படிப்பை முடிக்கலாம். ஆனால், காலப்போக்கில் கல்லூரி மாணவர்கள் பாடப்புத்தகங்கள் வாங்குவது அவசியமாகக் கருதலாம். இப்படிப்பட்ட மாற்றங்கள் வர வாய்ப்பு இருப்பதால் அடிப்படை ஆண்டு மிக அருகில் இருப்பது நல்லது. அடிப்படை ஆண்டு சாதாரண ஆண்டாக இல்லாமல் இருந்தால் அதற்குப் பக்கத்தில் உள்ள இரண்டு, மூன்று நான்கு அல்லது ஐந்து ஆண்டுகளையும் சேர்த்து அவற்றின் சராசரியை அடிப்படை ஆண்டுக்கு உரியதாகக் கொள்ளலாம். அதுபோலவே நடப்பு ஆண்டுக்கும் செய்யலாம்.

அட்டவணை 88இல் எடையிட்ட எளிய கூட்டுச் சராசரியைப் பயன்படுத்தி வாழ்க்கைச் செலவு குறியீட்டெண் பெறும் முறை தரப்படுகின்றது.

அட்டவணை - 88

பிரிவு	2010ஆம் ஆண்டிற்குக் குறியீட்டெண் (I)	எடை (W)	WI
உணவு	152	48	7296
எரிபொருள்	110	5	550
உடை	130	10	1300
வீட்டு வாடகை	100	12	1200
மற்றவை	80	15	1200
	மொத்தம்	**90**	**11546**

$$\text{வாழ்க்கைச் செலவுக் குறியீட்டெண்} = \frac{11546}{90} = 128.3$$

அடிப்படை ஆண்டை மாற்றும் முறை (Shifting base year)

இருவகையான சந்தர்ப்பங்களில் குறியீட்டெண்ணின் அடிப்படை ஆண்டை மாற்ற வேண்டிய அவசியம் வரலாம். இரண்டு நாடுகளில் விலைகள் எவ்வாறு கூடி அல்லது குறைந்துள்ளன எனப் பார்க்க வேண்டும். ஆனால், அந்த நாடுகளின் குறியீட்டெண்கள் வெவ்வேறு ஆண்டினை அடிப்படையாகக் கொண்டுள்ளன. இப்படிப்பட்ட சூழ்நிலையில் ஒரே மாதிரி அடிப்படை ஆண்டாகக் (உதாரணமாக 2000) கொண்டால் மட்டுமே ஒப்புமைப்படுத்துவது சரியாகும். இந்த நோக்கத்துடன் அட்டவணை 89 அமைக்கப் பெற்றுள்ளது.

அட்டவணை 89இல் முதலில் 2000ஆவது ஆண்டினை அடிப்படையாகக் கொண்டு கணிக்கப்பெற்ற குறியீட்டெண்கள் 2ஆவது நிரலில் தரப்பட்டுள்ளன; நிரல் 3இல், 2003ஆவது ஆண்டை அடிப்படையாகக் கொண்டு கணிக்கப்பெற்ற குறியீட்டெண்கள் தரப்பட்டுள்ளன.

அட்டவணை - 89

ஆண்டு	குறியீட்டெண்கள் 2000 = 100	புதிய குறியீட்டெண்கள் (2003 = 100)
2000	100	$\frac{100}{250} \times 100 = 40$
2001	110	$\frac{110}{250} \times 100 = 44$
2002	175	$\frac{175}{250} \times 100 = 70$
2003	250	$\frac{250}{250} \times 100 = 100$
2004	300	$\frac{300}{250} \times 100 = 120$
2005	400	$\frac{400}{250} \times 100 = 160$

சில சமயங்களில் அடிப்படை ஆண்டு (இங்கு 2000) மிகப் பழைய ஆண்டாக ஒப்புநோக்குவதற்குப் பொருந்தாத ஆண்டாக அமைந்திருக்கலாம். இந்நிலையில் புதிய ஆண்டு அல்லது பக்கத்தில் உள்ள ஆண்டு (இங்கு 2003) ஒப்புநோக்குவதற்குப் பொருத்தமான ஆண்டாக அமையலாம்.

இச்சூழ்நிலையில் புதிய ஆண்டின் குறியீட்டெண்ணை அடிப்படையாகக் (2003=100) கொண்டு புதிய குறியீட்டெண்ணைக் கண்டுபிடிக்கலாம்.

குறியீட்டெண்களில் விலைமாற்ற விளைவை நீக்கும் முறை (Deflating Index Numbers)

பொதுவாக எங்கும் விலையேற்றமே காணப்படுகிறது. இப்படிப்பட்ட சூழ்நிலையில் ஒருவரின் வருமானம் கூடியிருந்தால் அது விலை உயர்வினால் ஏற்பட்ட மாற்றமாகக்

கூட இருக்கலாம். உதாரணமாக, விலைகள் நான்கு மடங்குகள் (100லிருந்து 400 ஆக) கூடி, ஒருவருடைய பண வருமானம் இரண்டு மடங்கு மட்டுமே கூடியிருந்தால், உண்மையிலேயே அவரின் வருமானம் குறைந்துள்ளது என்றே பொருள். அந்நிலைகளில், கொடுக்கப்பட்டுள்ள மதிப்புக்களை, உரிய விலைக் குறியீட்டெண்களால் வகுப்பதன் மூலம் விலையேற்றத்தால் ஏற்பட்டிருக்கக்கூடிய மாற்றங்களை நீக்கிவிட முடியும். இவ்வாறாகவே, விலை இறக்கத்தால் ஏற்படக்கூடிய மாற்றங்களையும் நீக்கிவிட முடியும்.

உதாரணமாக, ஒருவருடைய ஒருநாள் வருமானம் 2000இல் 50 ரூபாயாகவும், 2009இல் 100 ரூபாயாகவும் இருந்ததெனக் கொள்வோம். அதேகாலகட்டத்தில் விலைக் குறியீட்டெண் 100ஆக இருந்து 125ஆகக் கூடியுள்ளது என்றும் கொள்வோம். அப்படியானால் 2009ஆம் ஆண்டு அவருடைய உண்மையான அல்லது சரிப்படுத்தப்பட்ட வருமானம் $\frac{100}{125} \times 100 =$ ரூ.80. எனவே, இவருடைய உண்மையான வருமானம் 30 ரூபாய் அளவுக்குக் கூடியுள்ளதே ஒழிய 50 ரூபாய் அல்ல. இந்த சரிசெய்யப்பட்ட வருமானம் (ரூ.80) விலைமாற்ற விளைவு நீக்கப்பட்ட வருமானம் (deflated or real income) என அழைக்கப்படுகிறது.

சிறந்த குறியீட்டெண்ணுக்குரிய சோதனைகள்

கீழ்க்காணும் முதல் மூன்று சோதனைகளை நிறைவு செய்யும் குறியீட்டெண் பிழையற்றதாக இருக்கும். அது விழுமிய குறியீட்டெண் (ideal or perfect) என அழைக்கப் பெறுகிறது.

(1) பொருள் திருப்புச் சோதனை (Commodity reversal test)

(2) காலத் திருப்புச் சோதனை (Time reversal test)

(3) பகுதி திருப்புச் சோதனை (Factor reversal test)

பொருள் திருப்புச் சோதனை (Commodity reversal test)

கொடுக்கப்பட்டுள்ள பொருட்களை எந்த வரிசையில் அடுக்கி (உதாரணமாக A, B, C, D or ACDB etc) குறியீட்டெண் கண்டுபிடித்தாலும் குறியீட்டெண்ணின் மதிப்பு மாறாமல் இருத்தல் வேண்டும். இச்சோதனையை எல்லாவகைக் குறியீட்டெண்களும் பூர்த்தி செய்கின்றன.

காலத் திருப்புச் சோதனை (Time reversal test)

மற்றவைகளை மாற்றாமல் வைத்து, அடிப்படை ஆண்டை நடப்பு ஆண்டாகவும், நடப்பு ஆண்டை அடிப்படை ஆண்டாகவும் மாற்றி குறியீட்டெண் கணக்கிடப்படும்போது, குறியீட்டெண்ணின் மதிப்பு மாறாமல் இருத்தல் வேண்டும். அதாவது விலைக்குறியீட்டெண் என்றால், $P_{01} \times P_{10} = 1$ எனவும், அளவுக் குறியீட்டெண் என்றால், $Q_{01} \times Q_{10} = 1$ எனவும் இருத்தல் வேண்டும்.

இச்சோதனையை நிறைவு செய்யும் குறியீட்டெண்கள் இரண்டு, ஒன்று ஃபிஷரின் குறியீட்டெண் மற்றொன்று மார்ஷல்-எட்ஜ்ஒர்த் குறியீட்டெண்.

பிஷரின் குறியீட்டெண் : $P_{01} = \sqrt{\frac{\Sigma P_1 Q_0}{\Sigma P_0 Q_0} \times \frac{\Sigma P_1 Q_1}{\Sigma P_0 Q_1}}$

இந்த சூத்திரத்தில் அடிப்படை ஆண்டை நடப்பு ஆண்டாகவும், நடப்பு ஆண்டை அடிப்படை ஆண்டாகவும் மாற்றினால்,

$$P_{10} = \sqrt{\frac{\Sigma P_0 Q_1}{\Sigma P_1 Q_1} \times \frac{\Sigma P_0 Q_0}{\Sigma P_1 Q_0}}$$

இவற்றிலிருந்து,

$$P_{01} \times P_{10} = \sqrt{\frac{\Sigma P_1 Q_0}{\Sigma P_0 Q_0} \times \frac{\Sigma P_1 Q_1}{\Sigma P_0 Q_1}} \times \sqrt{\frac{\Sigma P_0 Q_1}{\Sigma P_1 Q_1} \times \frac{\Sigma P_0 Q_0}{\Sigma P_1 Q_0}} = 1$$

இப்பொழுது மார்ஷல்- எட்ஜ்வொர்த்தின் குறியீட்டெண் எவ்வாறு காலத்திருப்புச்சோதனையைப் பூர்த்தி செய்கிறது எனப் பார்க்கலாம்.

மார்ஷல் - எட்ஜ்வொர்தின் விலைக்குறியீட்டெண்.

$$P_{01} = \frac{\Sigma (Q_0 + Q_1)P_1}{\Sigma (Q_0 + Q_1)P_0}$$

$$P_{10} = \frac{\Sigma (Q_1 + Q_0)P_0}{\Sigma (Q_1 + Q_0)P_1}$$

$$P_{01} \times P_{10} = \frac{\Sigma (Q_0 + Q_1)P_1}{\Sigma (Q_0 + Q_1)P_0} \times \frac{\Sigma (Q_1 + Q_0)P_0}{\Sigma (Q_1 + Q_0)P_1} = 1$$

லாஸ்பியர் மற்றும் பாஷேயின் குறியீட்டெண்கள் காலத்திருப்புச் சோதனையைப் பூர்த்தி செய்யவில்லை என்பது கீழே விளக்கப்படுகிறது.

(லாஸ்பியர்) $$P_{01} = \frac{\Sigma P_1 Q_0}{\Sigma P_0 Q_0}$$

$$P_{10} = \frac{\Sigma P_0 Q_1}{\Sigma P_1 Q_1}$$

$$P_{01} \times P_{10} = \frac{\Sigma P_1 Q_0}{\Sigma P_0 Q_0} \times \frac{\Sigma P_0 Q_1}{\Sigma P_1 Q_1} \neq 1$$

(பாஷே) $$P_{01} = \frac{\Sigma P_1 Q_1}{\Sigma P_0 Q_1}$$

$$P_{10} = \frac{\Sigma P_0 Q_0}{\Sigma P_1 Q_0}$$

$$P_{01} \times P_{10} = \frac{\Sigma P_1 Q_1}{\Sigma P_0 Q_1} \times \frac{\Sigma P_0 Q_0}{\Sigma P_1 Q_0} \neq 1$$

பகுதி திருப்புச் சோதனை (Factor reversal test)

இச்சோதனையின்படி, விலைக்குறியீட்டெண்ணில் Pயை Qஆகவும் Qவை Pஆகவும் மாற்றி எழுதினால், அளவுக் குறியீட்டெண் கிடைத்தல் வேண்டும். மேலும், விலைக் குறியீட்டெண்ணையும் அளவுக்குறியீட்டெண்ணையும் பெருக்கினால், மதிப்பு விகிதமாகிய,

$$\frac{\Sigma Q_1P_1}{\Sigma Q_0P_0}$$ கிடைத்தல் வேண்டும்.

இச்சோதனையை ஃபிஷரின் குறியீட்டெண் நிறைவு செய்கிறது. லாஸ்பியர், பாஷே, மார்ஷல்-எட்ஜ்வொர்த் ஆகியோரின் குறியீட்டெண்கள் இச்சோதனையைப் பூர்த்தி செய்யவில்லை.

ஃபிஷரின் விலைக்குறியீட்டெண் $= \sqrt{\frac{\Sigma P_1Q_0}{\Sigma P_0Q_0} \times \frac{\Sigma P_1Q_1}{\Sigma P_0Q_1}}$

இதனில், Pயை Qஆகவும் Qவை Pஆகவும் மாற்றினால்,

$$Q_{01} = \sqrt{\frac{\Sigma Q_1P_0}{\Sigma Q_0P_0} \times \frac{\Sigma Q_1P_1}{\Sigma Q_0P_1}}$$

இப்பொழுது

$$P_{01} \times Q_{01} = \sqrt{\frac{\Sigma P_1Q_0}{\Sigma P_0Q_0} \times \frac{\Sigma P_1Q_1}{\Sigma P_0Q_1}} \times \sqrt{\frac{\Sigma Q_1P_0}{\Sigma Q_0P_0} \times \frac{\Sigma Q_1P_1}{\Sigma Q_0P_1}}$$

$$= \sqrt{\frac{\Sigma P_1Q_1}{\Sigma P_0Q_0}} \times \sqrt{\frac{\Sigma Q_1P_1}{\Sigma Q_0P_0}}$$

$$= \frac{\Sigma P_1Q_1}{\Sigma P_0Q_0}$$ = மதிப்பு விகிதம்.

வட்டமான சோதனை (Circular test or cyclicai test)

காலத்திருப்புச்சோதனையின் நீட்சியே வட்டமான சோதனை ஆகும்.

ஒரு குறியீட்டெண் கீழ்க்கண்ட நிபந்தனைக்குட்பட்டு இருக்கும்போது வட்டமான சோதனை பூர்த்தி செய்யப்படுவதாக கூறப்பெறுகிறது.

$$P_{12} \times P_{23} \times P_{34} \times \ldots P_{(n-1)n} \times Pn1 = 1$$

Pn1 ஆனது nஆவது ஆண்டை அடிப்படையாகக் கொண்டு முதலாவது ஆண்டிற்குக் கணிக்கப்பெற்ற குறியீட்டெண்.

ஒரே ஒரு பொருள் மட்டும் தரப்பெற்றிருக்கும்பொழுது இச்சோதனை பூர்த்தியாகிறது. ஃபிஷரின் குறியீட்டெண் இச்சோதனையைத் தோராயமாகப் பூர்த்தி செய்கிறது. ஃபிஷரின் குறியீட்டெண்ணை இரு குறிப்பிட்ட வருடங்களில் உள்ள மதிப்புக்களை ஒப்பிடுவதற்கு மட்டும் பயன்படுத்தாததாலும், பல ஆண்டுகளின் மதிப்புக்களை ஒப்பிடுவதற்குப் பயன்படுத்துவதாலும், இச்சோதனையை ஃபிஷரின் குறியீட்டெண் பூர்த்தி செய்யாவிடில் பரவாயில்லை எனக் கூறப்படுகிறது.

விகிதாச்சாரச் சோதனை (Proportionality test)

எல்லாப் பொருள்களின் விலைகளும் சமமாக ஒரே விகிதத்தில் (λ) கூடினால், விலைக்குறியீட்டெண் λக்குச் சமமாக இருக்கும்.

மேலே கொடுக்கப்பட்டுள்ள ஐந்து சோதனைகளையும் திருப்திப்படுத்தும் ஏதேனும் ஒரு குறியீட்டெண் உள்ளதா எனும் கேள்வி இங்கு எழலாம். ஒரே ஒரு பொருளும் அதன் விலையும் மட்டும் இருந்தால், விலை மற்றும் அளவு (quantity) சார்பிகள் இந்த ஐந்து சோதனைகளையும் நிறைவு செய்யும். ஒன்றுக்கு மேற்பட்ட பொருளும் விலையும்

இருந்தால், எந்தக் குறியீட்டெண்ணும் ஐந்து சோதனைகளையும் நிறைவு செய்யமாட்டாது. ஃபிஷர் உடைய குறியீட்டெண் வட்டமான சோதனையைத் தவிர மற்ற நான்கு சோதனைகளை மட்டும்தான் நிறைவு செய்கிறது.

மேலே கூறப்பட்டுள்ள ஐந்து சோதனைகளையும் நிறைவு செய்வது பற்றி விளக்குவதற்கு விலையும் அளவும் (quantity) தேவைச் சார்பு (demand function) மூலம் தொடர்புபடுத்தப்பட்டுள்ள உண்மையை அறிய வேண்டி உள்ளது.

முன்னுரிமைச் சார்பு (preference function) சமபடித்தானதாக (homothetic) இருந்தால்தான், மேலே கூறப்பெற்றுள்ள ஐந்து சோதனைகளையும் நிறைவு செய்யும் ஒரு குறியீட்டெண் இருக்க முடியும் என நாகரும் தாஸும் (NAGAR A.L. & DAS R.K.) தங்கள் புத்தகத்தில் (Basic Statistics, Second Edition, Oxford University Press, Delhi, 1983) 306ஆம் பக்கத்தில் குறிப்பிட்டுள்ளார்கள்.

நடைமுறை குறியீட்டெண்கள் (Official Index Numbers)

மேலே சில வகைக் குறியீட்டெண்கள் பற்றிய விளக்கங்கள் தரப்பெற்றன. நடைமுறையில் குறியீட்டெண்கள் தயாரிப்பது அவ்வளவு எளிதல்ல. ஒரு நடவடிக்கை சம்மந்தமாக குறியீட்டெண் தயாரிப்பதற்கு அது தொடர்பான பல பொருள்கள் பற்றிய அளவுகளையும் விலைகளையும் பற்றிய விபரங்கள் தேவைப்படுகின்றன. உதாரணமாக, மொத்த விலைக் குறியீட்டெண் தாயரிப்பதற்கு, நம் நாட்டில் (இந்தியாவில்) மூன்று வகைகளாகப் பொருள்களைப் பிரித்தனர். அவை : (1) அடிப்படைப் பண்டங்கள் (Primary articles), (2) எரிசக்திகள் (Fuel, power, light) மற்றும் உயவுப் பொருள்கள் (Lubricants) (3) உற்பத்தி செய்யப்பட்ட தொழில் பொருள்கள் (Manufactured products). அடிப்படைப்

பண்டங்களில் மேலும் மூன்று பொருட்களும், இரண்டாவது வகையில் மூன்று பொருட்களும் மூன்றாவது வகையில் 11 வகைப் பொருட்களுமாக மொத்தம் 17 பொருட்கள் பற்றிய விபரங்கள் சேகரிக்கப்படுகின்றன. அடிப்படைப் பண்டங்களுக்குள், 39 வகையான உணவுப் பண்டங்களும் 26 வகையான உணவல்லாத பொருள்களும் 15 தாதுப் பொருள்களுமாக 80 வகையான பொருள்கள் உள்ளன. இவ்வாறாக, மொத்தம் 360 பொருள்களுக்கான விபரங்கள் சேகரிக்கப்படுகின்றன. இந்த 360 பொருள்களுக்கும் விலைக்குறிகள் (price quotations) தேர்ந்தெடுக்கப்பட்ட சந்தைகளில் ஒவ்வொரு வெள்ளிக்கிழமையும் சேகரிக்கப் படுகின்றன. இவ்வாறாக 1275 விலைக்குறிகள் எடுக்கின்றனர். இந்த விலை, தீர்வைவரி (excise duty) சேர்க்கப்பட்ட உற்பத்தியாளர்கள் விலையாகும் (producers' price). இவற்றின் மொத்த எடை (weights) 1000 ஆகும்.

அதுபோல விவசாய உற்பத்திக்கான குறியீட்டெண்களும் (Index Numbers of Agricultural Production) தயாரிக்கப் படுகின்றன. அதற்கு தானியங்கள், பருப்பு வகைகள், நார்ச்சத்துப் பொருட்கள் (fibres), தோட்டப் பொருட்கள் (Plantations), காரம் மற்றும் வாசனைப் பொருட்கள் (condiments and spices), காய்கறிகள் மற்றும் பழங்கள் இன்னும் இதர விளைபயிர்கள் ஆகியவற்றின் விபரங்கள் சேகரிக்கப்படுகின்றன.

அத்துடன் விவசாயக் கூலியாட்களின் நுகர்வு விலைக் குறியீடெண்ணும் தயாரிக்கப் பெறுகிறது.

இதுவரை கூறப் பெற்றுள்ள மூன்று வகைக் குறியீட்டெண்கள் பற்றி மேலும் விபரங்கள் அறிய விரும்புபவர்கள், நாகர் மற்றும் தாஸ் எழுதிய (முன்னர் கூறப்பட்டுள்ள) புத்தகத்தில் 293ஆம் பக்கம் முதல் 301ஆம் பக்கம் வரை பார்க்கலாம்.

உண்மை வாழ்க்கைச் செலவு குறியீட்டெண் (True Cost of Living Index)

உண்மை வாழ்க்கைச் செலவுக் குறியீட்டெண் என்பது இரண்டுவிதமான விலைகள் நிலவும் சூழ்நிலையில் ஒரு குறிப்பிட்ட அளவு பயன்பாட்டைப் பெறுவதற்குத் தேவைப்படும் மிகச்சிறு செலவுகளின் விகிதாச்சாரம் ஆகும் (Nagar, A.L., & Das, R.K, 1983, Basic Statistics, Second Edition, Oxford University Press, Delhi, p. 302)

இதனைச் சுருக்கமாக $P^{o}_{01} \equiv P_{01}\,[u(q_0)] = \dfrac{E\,(p_1/q_0)}{E\,(p_0/q_0)}$

இதில் அடிப்படை ஆண்டில் பெற்ற பயன்பாடு (utility) குறிப்பிட்ட பயன்பாடு அளவாகக் கொள்ளப் பெற்றுள்ளது. நடப்பு ஆண்டில் பெற்ற பயன்பாட்டினை (utility) குறிப்பிட்ட பயன்பாடு அளவாகக் கொண்டால், உண்மை வாழ்க்கைச் செலவுக் குறியீட்டெண்ணை

$$p^{1}_{01} \equiv p_{01}\,[u(q_1)] = \frac{E\,(p_1/q_1)}{E\,(p_0/q_1)}$$ எனலாம்.

p^{o}_{01}, p^{1}_{01}க்கும் லாஸ்பியர் மற்றும் பாஷே விலைக் குறியீட்டெண்களுக்கும் நெருங்கிய தொடர்புகள் உள்ளன. அதாவது,

$$p^{0}_{01} \leq p^{L}_{01}$$

இதில் p^{L}_{01} என்பது லாஸ்பியரின் விலைக்குறியீட்டெண். அதுபோல,

$$p^{1}_{01} \geq p^{P}_{01}$$

இதில் p^{P}_{01} என்பது பாஷேயின் விலைக்குறியீட்டெண்.

சமபடித்தான பயன்பாட்டுச் சார்பில் (homothetic utility function), ஆரம்பக்கட்ட பயன்பாட்டு அளவு (reference utility

level) அடிப்படை ஆண்டாக இருந்தாலும் சரி, நடப்பு ஆண்டாக இருந்தாலும் சரி, கீழ்வரும் தொடர்பு உண்மையாக இருக்கும்.

$$p^{P}_{01} \leq p_{01} \leq p^{L}_{01}$$

இதில் p_{01} என்பது உண்மை வாழ்க்கைச் செலவுக் குறியீட்டெண். சமபடித்தான பயன்பாட்டுச் சார்பு இல்லாதபோது p_{01},p^{L}_{01}க்குக் குறைவாகவும் p^{P}_{01}க்கு அதிகமாகவும் இருக்கும் என்று சொல்ல முடியாது.

இந்தத் தொடர்புகளைத் தெளிவாகப் புரிந்து கொள்ள நாகர் மற்றும் தாஸ் ஆகியோரின் புத்தகத்தில் (Basic Statistics, Second Edition), 301ஆம் பக்கத்தில் இருந்து 305ஆம் பக்கம் வரைப் படிக்கலாம். இந்தத் தொடர்புகளை விளக்குவதற்கு அவர்கள் பயன்பாட்டுச் சமநோக்கு வளைகோடுகளைப் (indifference curve) பயன்படுத்தியுள்ளனர். இதன் மூலம் அவர்கள், புள்ளியியல் நுண்மைப் பொருளியலை (microeconomics) நன்றாகப் புரிந்து கொள்ளவும், நுண்மைப் பொருளியல், புள்ளியியலை நன்றாகப் புரிந்து கொள்ளவும் உதவுகின்றன என்னும் கருத்தினையும் வலியுறுத்தி உள்ளனர். அதுமட்டுமல்லாது, எல்லாப் பாடங்களையும் ஒன்றோடொன்று தொடர்புப்படுத்தி படிப்பது புரிந்து கொள்ளுதலை எளிதாகவும், சுவையானதாகவும் நீண்ட நாட்கள் நினைவில் வைத்துக்கொள்ள உதவும் என்ற நம்பிக்கையையும் வெளிப்படுத்தி உள்ளார்கள்.

ஃ

கூடுதல் பயிற்சிகள்

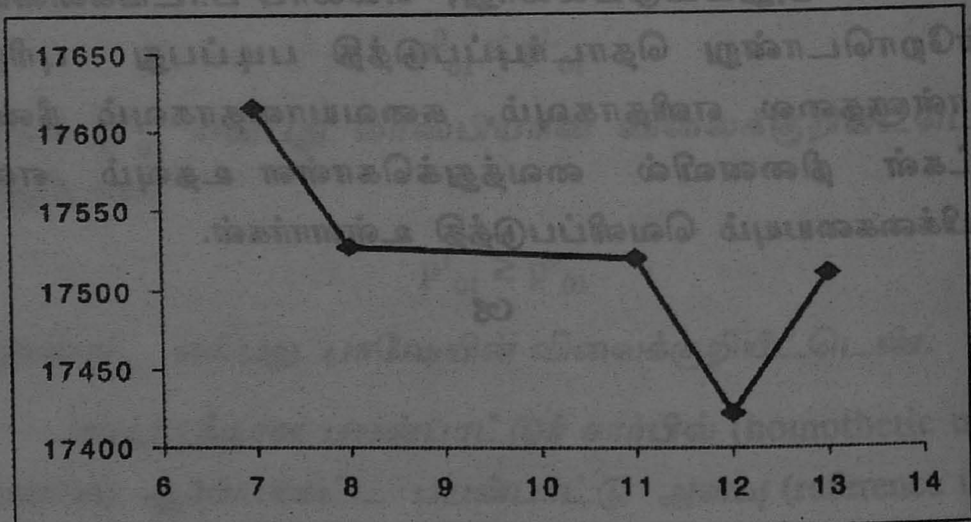
ADDITIONAL EXERCISES

1. கீழ்வரும் விபரங்களைச் சரியான பொருத்தமான முறையைப் பயன்படுத்திக் காண்பிக்கவும்.

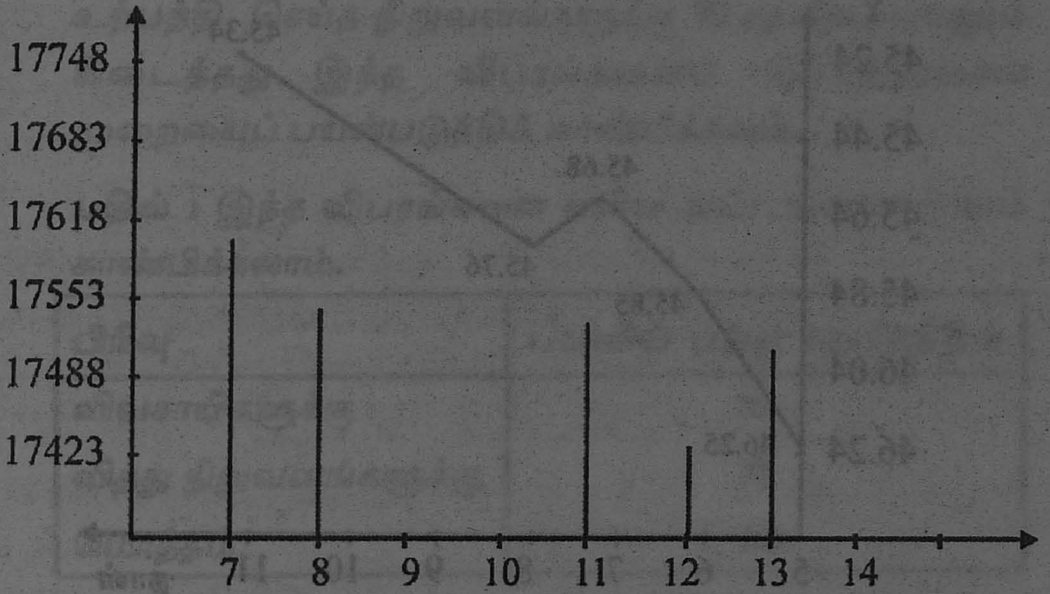
ஆண்டு	மாதம்	நாள்	பாம்பே பங்குச் சந்தையின் உணர்வுக் குறியீட்டெண் (BSE SENSEX)
2010	ஜனவரி	7	17615.72
2010	ஜனவரி	8	17526.71
2010	ஜனவரி	11	17519.00
2010	ஜனவரி	12	17422.51
2010	ஜனவரி	13	17509.80

பதில் : இங்கு கொடுக்கப்பட்டுள்ள மாறி தொடர் மாறியாக (Continuous variable) இருப்பதால் வரைபடம் (graph) மிகப் பொருத்தமாக இருக்கும். கோட்டு விளக்கப்படமும் (line diagram) வரையலாம்.

வரைபடம்
BSE SENSEX



கோட்டு விளக்கப்படம்



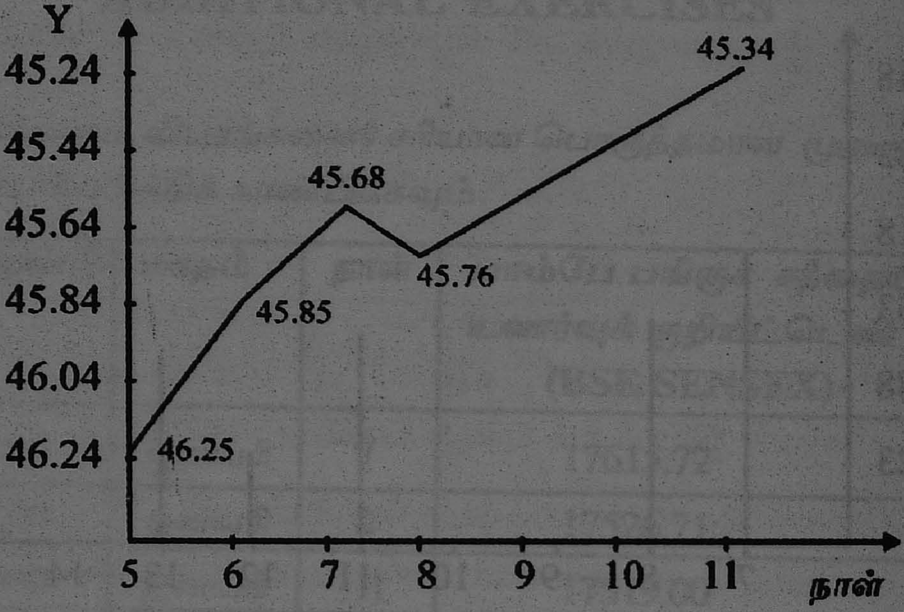
2. கீழ்வரும் விபரங்களைச் சரியான பொருத்தமான முறையைப் பயன்படுத்தி காண்பிக்கவும்.

ஆண்டு	மாதம்	நாள்	இந்திய ரூபாயில் அமெரிக்காவின் 1 டாலரின் மதிப்பு
2010	ஜனவரி	5	46.25
2010	ஜனவரி	6	45.85
2010	ஜனவரி	7	45.68
2010	ஜனவரி	8	45.76
2010	ஜனவரி	11	45.34

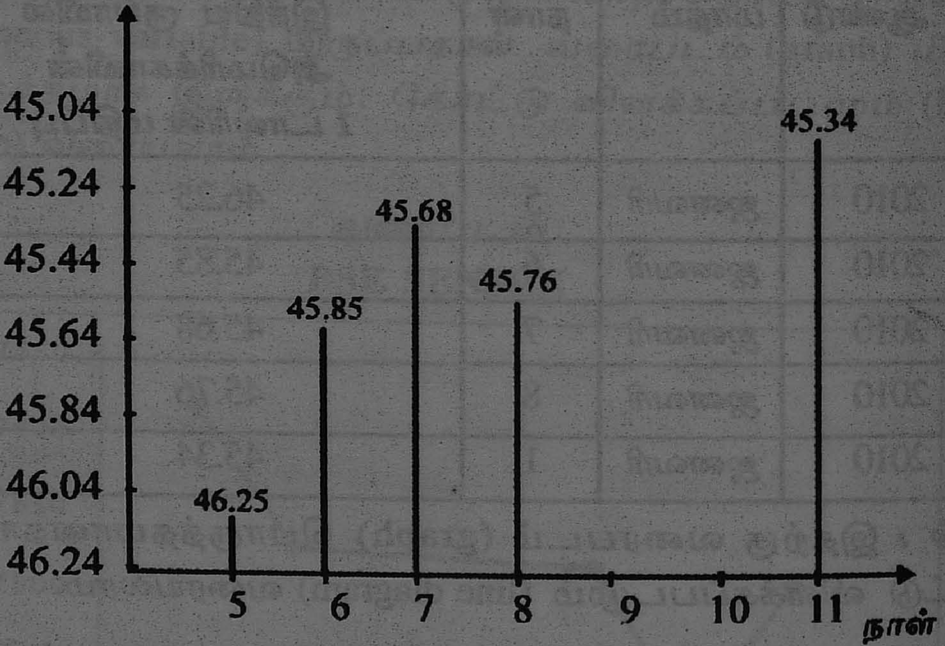
பதில் : இதற்கு வரைபடம் (graph) பொருத்தமானதாகும். கோட்டு விளக்கப்படமும் (line diagram) வரையலாம்.

இங்கு இந்தியப் பணத்தின் மதிப்பு உயரும்போது கோடு மேலே செல்வது பொருத்தமாகும். எனவே, செங்குத்து (y) அச்சில் எண்கள் மாற்றி எழுதப்படுகின்றன.

வரைபடம்



கோட்டு விளக்கப்படம்

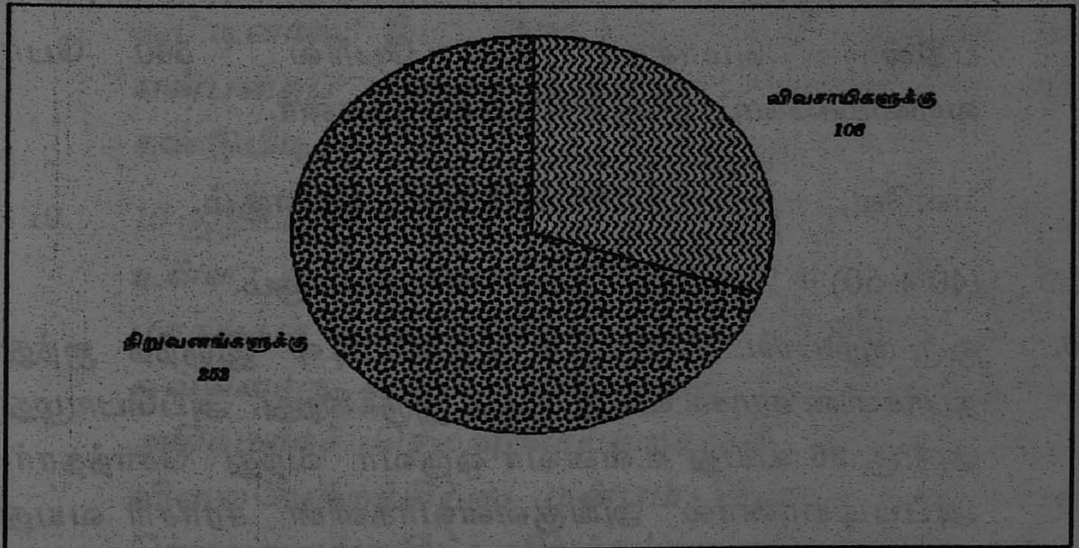


3. உயர்ரக வித்துக்களைப் பயன்படுத்தியதால் விவசாயிகளுக்கு 30 சதவீதப் பயனும் வித்துக்களை உற்பத்தி செய்த நிறுவனங்களுக்கு 70 சதவீதப் பயனும் கிடைத்தது. இந்த விபரங்களைப் பொருத்தமான முறையைப் பயன்படுத்திக் காண்பிக்கவும்.

பதில் : இந்த விபரங்களை எளிய அட்டவணை மூலம் காண்பிக்கலாம்.

பிரிவு	பயனின் பங்கு சதவீதத்தில்
விவசாயிகளுக்கு	30
வித்து நிறுவனங்களுக்கு	70
மொத்தம்	100

இந்த விபரங்களை வட்ட விளக்கப்படம் மூலமும் காண்பிக்கலாம். வட்டத்தில் மொத்தம் 360 பாகை உள்ளதால் 100 சதவீதத்தை 360 பாகைக்குச் சமமாக ஆக்க வேண்டும். அப்படியானால் 1 சதவீதத்திற்கு 3.60 பாகை. எனவே, விவசாயிகளுக்கு 108 ; நிறுவனங்களுக்கு 252 பாகை; மொத்தம் 360 பாகை.



4. ஒருவர் 10 மாணவர்களின் சராசரி மதிப்பெண் 65 என்று கூறினார். பிறகு ஒவ்வொரு மாணவரின் மதிப்பெண்ணையும் பார்த்தபோது, ஒரு மாணவரின் மதிப்பெண் 13க்குப் பதிலாக 30 என்று சேர்க்கப்பட்டு இருந்தது. அப்படியானால் உண்மையான சராசரி எவ்வளவு?

பதில் : 10 மாணவர்களின் மொத்த மதிப்பெண் 650. இதில் தவறான மதிப்பெண்ணைக் கழிக்க வேண்டும். சரியான மதிப்பெண்ணைக் கூட்ட வேண்டும். இவ்வாறு செய்த பின்னர் உண்மையான மொத்த மதிப்பெண் 650 - 30 + 13 = 633. **எனவே உண்மையான சராசரி 63.3 ஆகும்.**

5. 100 நபர்கள் கொண்ட ஒரு கிராமத்தில் 40 சதவீத நபர்கள் வறுமைக்கோட்டிற்குக் கீழே வாழ்ந்தனர். இன்னொரு கிராமத்தில் 1000 பேர் வாழ்ந்தனர். அங்கு 50 சதவீத நபர்கள் வறுமைக்கோட்டிற்குக் கீழே வாழ்ந்தனர். அப்படியானால், இரண்டு கிராமங்களிலும் சேர்த்து எத்தனை சதவீத நபர்கள் வறுமைக்கோட்டிற்குக் கீழே வாழ்ந்தனர்?

பதில் : மொத்தம் 1100 பேரில் 540 பேர் வறுமைக்கோட்டிற்கும் கீழே வாழ்ந்தனர்.

எனவே, $\frac{540}{1100} \times 100$ **சரியான பதிலாகும்.**

(40 + 50) ÷ 2 என்பது தவறான பதிலாகும்.

6. ஒரு குறிப்பிட்ட நாளில் ஒரு வீட்டில் இருந்த ஐந்து நபர்களின் சரா்சரி வயது 20ஆக இருக்கிறது. அப்பொழுது அங்கு 26 வயது உள்ளவர் ஒருவர் வந்து சேர்ந்தார். அப்படியானால் அங்குள்ளவர்களின் சராசரி வயது எவ்வளவு?

பதில் : (100 + 26) ÷ 6 = 21 வயது.

7. 10 குழந்தைகளின் சராசரி எடை ஒரு நாள் 6 கிலோவாக இருந்தது. சில நாட்கள் கழித்து, அந்த 10 குழந்தைகளில் ஒவ்வொரு குழந்தையின் எடையும் 500 கிராம் கூடியிருந்தது. இப்பொழுது, அந்தக் குழந்தைகளின் சராசரி எடை எவ்வளவு?

பதில் : 6.5 கிலோ

8. வெவ்வேறு வயது கொண்ட 10 மாணவர்களின் சராசரி வயது 20 ஆக இருந்தது. 5 ஆண்டுகள் கழித்து அவர்களின் சராசரி வயது எவ்வளவாக இருக்கும்?

பதில் : 25 வயது

9. 10 மாணவர்கள் பெற்ற மதிப்பெண்களின் திட்ட விலக்கம் 15ஆக இருந்தது. திட்ட விலக்கம் கணக்கிட்ட பின்னர் அவர்களின் ஆசிரியை ஒவ்வொரு மாணவருக்கும் 10 மதிப்பெண்கள் கூடுதலாக வழங்கினார். அப்படியானால் புதிய திட்டவிலக்கம் எவ்வளவு?

பதில் : 15

எல்லா மதிப்புகளுடனும் ஒரு மாறிலியைக் கூட்டினால், திட்டவிலக்கம் மாறாது. இது சரிதானா என்பதைப் பார்க்க அடுத்த கேள்விக்குப் பதில் கண்டுபிடிக்கலாம்.

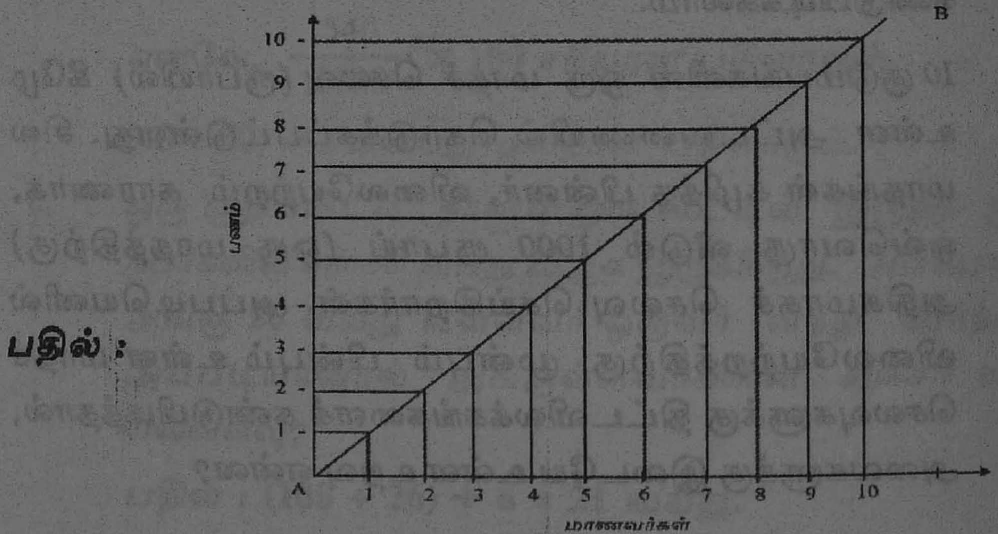
10. 10 குடும்பங்களின் ஒரு மாதச் செலவு (ரூபாயில்) கீழே உள்ள அட்டவணையில் கொடுக்கப்பட்டுள்ளது. சில மாதங்கள் கழித்த பின்னர், விலையேற்றம் காரணாக, ஒவ்வொரு வீடும் 1000 ரூபாய் (ஒரு மாதத்திற்கு) அதிகமாகச் செலவு செய்கிறார்கள். அப்படியெனில் விலையேற்றத்திற்கு முன்பும் பின்பும் உள்ள மாதச் செலவுகளுக்கு திட்டவிலக்கங்களைக் கண்டுபிடித்தால், அவைகளுக்கு இடையே உள்ள உறவு என்ன?

வீடு	மாதச் செலவு (ரூ)	விலையேற்றம் காரணமாக கூடிய தொகை
1	500	1000
2	600	1000
3	550	1000
4	450	1000
5	750	1000
6	700	1000
7	800	1000
8	560	1000
9	850	1000
10	900	1000

பதில் : விலையேற்றத்திற்கு முன்பும் பின்பும் உள்ள மாதச் செலவுகளுக்கு திட்டவிலக்கங்களைக் கண்டுபிடித்தால், அவை சமமாக இருக்க வேண்டும்.

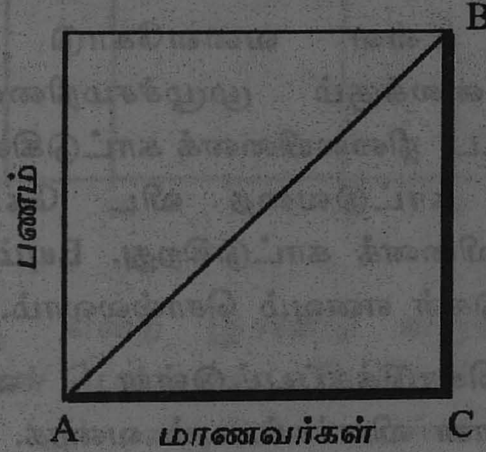
11. ஒரு வகுப்பில் 10 மாணவர்கள் இருக்கிறார்கள். ஒவ்வொரு மாணவரும் ரூபாய் 1 வைத்திருக்கிறார். மாணவர்களின் குவிவு அலைவெண்களை படுக்கை (x) அச்சிலும், பணத்தின் குவிவு அலைவெண்களை செங்குத்து (y) அச்சிலும் வைத்து ஒரு விளக்கப்படம் வரையவும்.

பதில் :



இதில் AB என்ற நேர்கோடு முழு சமநிலையைக் (Perfect equality) காட்டுகிறது. 5 மாணவர்களிடம் 5 ரூபாயும், 6 மாணவர்களிடம் 6 ரூபாயும் ஆகக் காட்டுகிறது.

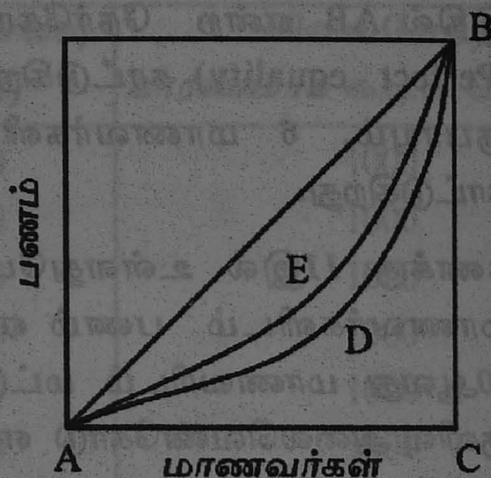
12. கணக்கு 11இல் உள்ளதுபோல் அல்லாது முதல் 9 மாணவர்களிடம் பணம் ஏதும் இல்லையென்றும், 10ஆவது மாணவரிடம் மட்டும் 10 ரூபாய் இருந்தால் குவிவு அலைவெண்கோடு எப்படி இருக்கும்?



பதில் : படத்தில் காண்பதுபோல ACB ஆக இருக்கும். முதல் ஒன்பது மாணவர்களிடம் பணம் இல்லை. எனவே அலைவெண்கோடு படுக்கையாக (0) உள்ளது. 10ஆவது மாணவரிடம் ரூ.10 இருப்பதால், கோடு Cயிலிந்து 'B'க்குத் தாவி விடுகிறது. இதில் ACB முழுச் சமநிலையின்மை (perfect inequality) அல்லது முழு ஏற்றத்தாழ்வினைக் காட்டுகிறது.

13. முழுச் சமநிலைக்கும் முழுச் சமநிலையின்மை அல்லது முழு ஏற்றத்தாழ்வுக்கும் இடைப்பட்ட நிலையினைக் காண்பிக்கவும்.

பதில் :



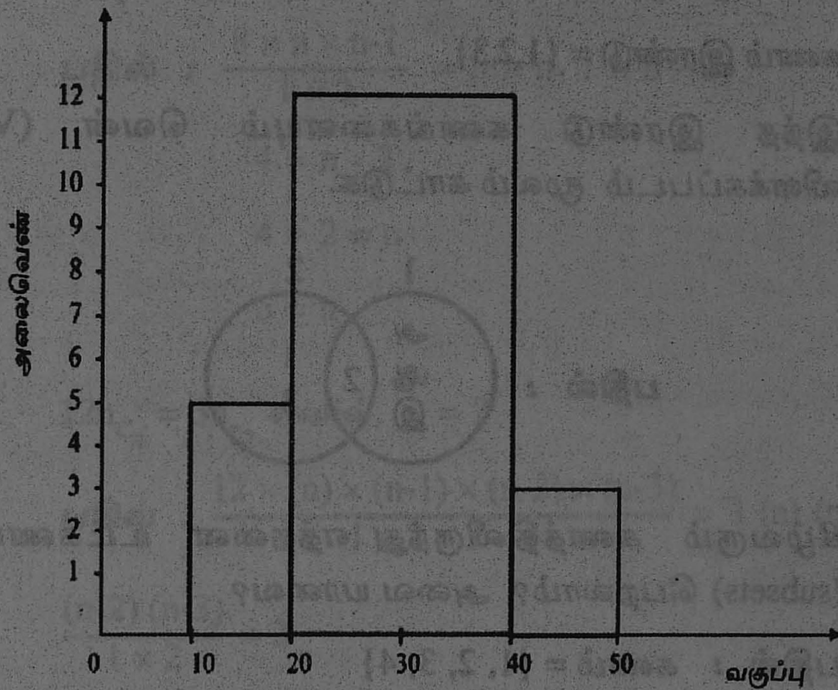
படத்தில் உள்ள வளைகோடு Eயும் Dயும் முழுச்சமநிலைக்கும் முழுச்சமநிலையின்மைக்கும் இடைப்பட்ட நிலையினைக் காட்டுகின்றன. அதிலும் கோடு E காட்டுவதை விட கோடு D அதிக ஏற்றத்தாழ்வினைக் காட்டுகிறது. Eயும் Dயும் லாரன்ஸ் வளைகோடுகள் எனவும் சொல்லலாம்.

14. கீழே கொடுக்கப்பட்டுள்ள விபரங்களுக்குப் பொருத்தமான விளக்கப்படம் வரைக.

வகுப்பு	அலைவெண்
10-19	5
20-39	12
40-49	3

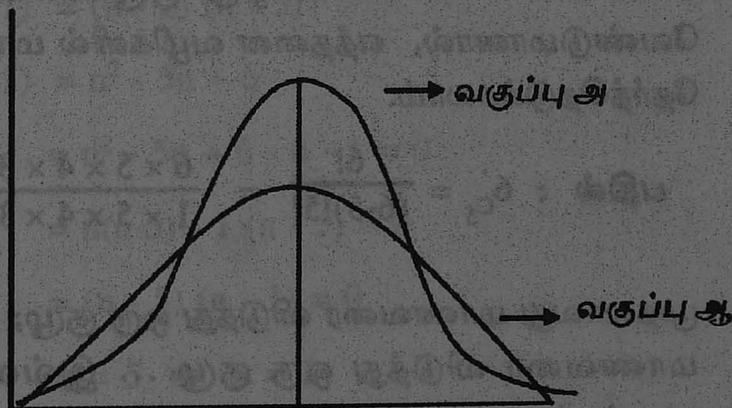
பதில் : கொடுக்கப்பட்டுள்ள வகுப்புகளில் இடைவெளிகள் வித்தியாசமாக இருப்பதால், இதற்கு பரவல் செவ்வகப் படம் பொருத்தமாக இருக்கும்.

வகுப்பு	அலைவெண்	அகலம்
10-19	5	10
20-39	12	20
40-49	3	10



15. 'அ', 'ஆ' என்ற இரண்டு வகுப்புகளில் உள்ள மாணவர்கள் பெற்றுள்ள மதிப்பெண்களின் சராசரி 70 எனவும், 'அ' வகுப்பில் உள்ள மாணவர்களின் மதிப்பெண்கள் சராசரியை ஒட்டிக் குவிந்தும், வகுப்பு 'ஆ'வில் உள்ள மாணவர்களின் மதிப்பெண்கள் மிகப் பரவலாகவும், அந்தப் பரவல்கள் சமச்சீரானவை எனவும் கொண்டு வரைபடம் ஒன்று வரைக.

பதில் :

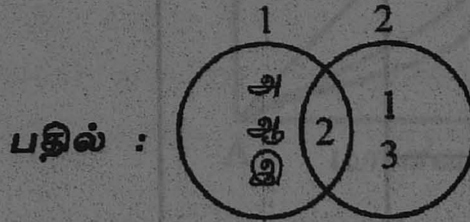


16. கணம் ஒன்று = {அ, ஆ, இ, 2}

கணம் இரண்டு = {1,2,3}

இந்த இரண்டு கணங்களையும் வென் (Venn) விளக்கப்படம் மூலம் காட்டுக.

பதில் :



17. கீழ்வரும் கணத்திலிருந்து எத்தனை உட்கணங்கள் (subsets) பெறலாம்? அவை யாவை?

பதில் : கணம் = {1, 2, 3, 4}

உட்கணங்களின் எண்ணிக்கை = $2^n = 2^4 = 16$. அவை

{}, {1}, {2}, {3}, {4}, {1,2}, {1,3}, {1,4}, {2,3}, {2,4}, {3,4}, {1,2,3}, {1,2,4}, {1,3,4}, {2,3,4}, {1,2,3,4}

18. X எனும் மாறியின் மாறுபாடு (variance) யாது?

பதில் : $\text{var}(X) = E[X - E(X)]^2 = E(X-\mu)^2$

19. ஒரு வகுப்பில் 6 மாணவர்கள் உள்ளனர். அதிலிருந்து 5 மாணவர்கள் கொண்ட ஒரு குழுவினை அமைக்க வேண்டுமானால், எத்தனை வழிகளில் மாணவர்களைத் தேர்ந்தெடுக்கலாம்.

பதில் : $6_{C_5} = \frac{6!}{(6-5)!5!} = \frac{6 \times 5 \times 4 \times 3 \times 2 \times 1}{1 \times 5 \times 4 \times 3 \times 2 \times 1} = 6$

முதலாவது மாணவரை விடுத்து ஒரு குழு; இரண்டாவது மாணவரை விடுத்து ஒரு குழு ... இவ்வாறாக ஆறு குழுக்கள்.

20. $8n_{C_2} = n_{P_3}$ எனில், n = ?

பதில் : $\frac{8 \times n \times n\text{-}1}{1 \times 2} = n \times n\text{-}1 \times n - 2$

$4 = n - 2$

$4 + 2 = n$

$6 = n$

20. $12n_{C_4} = 3n_{P_2}$ எனில், n = ?

பதில் : $\frac{12 \times (n) \times (n\text{-}1) \times (n\text{-}2) \times (n\text{-}3)}{1 \times 2 \times 3 \times 4} = 3\ (n)\ (n\text{-}1)$

$\frac{(n\text{-}2)\ (n\text{-}3)}{1 \times 2} = 3$

$(n - 2)\ (n\text{-}3) = 6$

$n^2 - 5n + 6 = 6$

$n^2 = 5n$

n = 5

21. $(n + 1)_{P_3} = n_{P_4}$ எனில், n = ?

பதில் : $(n + 1)\ (n)\ (n\text{-}1) = (n)\ (n\text{-}1)\ (n\text{-}2)\ (n\text{-}3)$

$(n + 1) = (n - 2)\ (n - 3)$

$(n + 1) = n^2 - 5n + 6$

$= n^2 - 5n + 6 - n - 1 = 0$

$= n(n\text{-}5) - 1\ (n - 5) = 0$

$= (n - 5)\ (n - 1) = 0$

n = 5 or n = 1

22. Xன் மதிப்பை 2 படிச் சமன்பாட்டில் கண்டுபிடிக்க பயன்படும் சூத்திரம் யாது?

$$X = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

இருபடிச் சமன்பாட்டில் உள்ள Xன் மதிப்பைக் காரணிப்படுத்தியும் (Factorisation) காணலாம். கணக்கு 20, 21 மற்றும் 23இல் காரணிப்படுத்தப்பட்டுள்ளதைப் பார்க்கலாம்.

23. $n_{p_2} = 20$எனில், n = ?

பதில் : $(n)(n-1) = 20$

$n^2 - n - 20 = 0$

$n^2 - 5n + 4n - 20 = 0$

$n(n-5) + 4(n-5) = 0$

$(n-5)(n+4) = 0$

$n = -4$ **or** $n = 5$

24. $3 \times {}^{n+1}c_3 = 7 \times n_{c_2}$ எனில், n = ?

பதில் : $\frac{3 \times (n+1)(n)(n-1)}{3 \times 2 \times 1} = \frac{7 \times (n)(n-1)}{2 \times 1}$

$\frac{3 \times (n+1)}{3 \times 2 \times 1} = \frac{7}{2}$

$\frac{(n+1)}{2 \times 1} = \frac{7}{2}$

$(n+1) = \frac{14}{2} = 7$

$n = 6$

25. 100 தடவைகள் ஒரு நல்ல நாணயத்தைச் சுண்டிப் போட்டதில் 56 தடவைகள் தலைகள் மேலே தெரிந்துள்ளன. அப்படியானால், 101ஆவது தடவை அந்த நாணயத்தைச் சுண்டினால் பூ வருவதற்குள்ள நிகழ்தகவு யாது?

பதில் : பூ வருவதற்கு நிகழ்தகவு இந்தச் சோதனையின்படி *44/100* ஆக இருப்பதால்,

101ஆவது முறை அந்த நாணயத்தைச் சுண்டும் போது பூ வருவதற்காந நிகழ்தகவு 0.44 ஆகும்.

26. ஒரு பெட்டியில் 6 சிவப்பு, 4 வெள்ளை, 5 கருப்புப் பந்துகள் உள்ளன. அப்பெட்டியிலிருந்து ஒன்றன்பின் ஒன்றாக 3 பந்துகள் எடுக்கப்பட்டால் அது சிவப்பு (Red), வெள்ளை (W) கருப்புப் (B) பந்தாக இருப்பதற்கு நிகழ்தகவு யாது? (அ) எடுத்த பந்தை மீண்டும் பெட்டியில் போட்டால் (ஆ) எடுத்த பந்தை பெட்டியில் போடாவிட்டால்?

பதில் : (அ) எடுத்த பந்தை மீண்டும் பெட்டியில் போட்டு எடுத்தால்,

$$\left(\frac{6}{6+4+5}\right)\left(\frac{4}{6+4+5}\right)\left(\frac{5}{6+4+5}\right) = \frac{8}{225}$$

(ஆ) எடுத்த பந்தை பெட்டியில் போடாமல் எடுத்தால்,

$$\left(\frac{6}{6+4+5}\right)\left(\frac{4}{5+4+5}\right)\left(\frac{5}{5+3+5}\right) = \frac{4}{91}$$

27. 0 முதல் 9 வரையுள்ள 10 எண்களைக் கொண்டு எத்தனை நான்கு இலக்க எண்கள் எழுதலாம்? (அ) வந்த எண்கள் திரும்ப வரலாம் என்றால் (ஆ) வந்த எண்கள் திரும்ப வரக்கூடாது என்றால் (இ) வந்த எண்கள் திரும்ப

வரலாம், ஆனால் கடைசி இலக்கம் பூஜ்யமாக இருக்க வேண்டும் என்றால்?

பதில் :

(அ) முதல் இலக்கம் பூஜ்யம் வரக்கூடாதென்பதால்

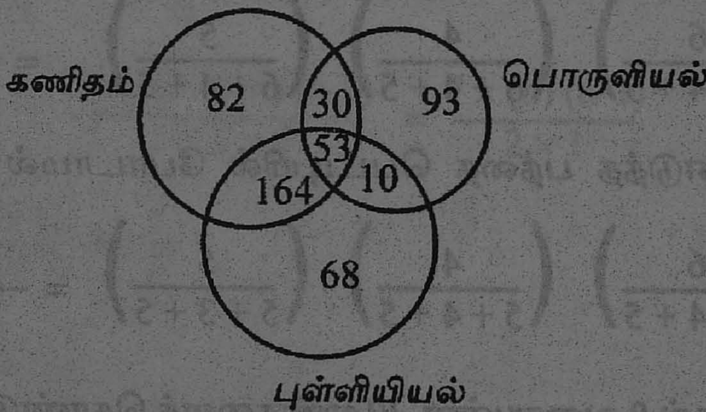
$9 \times 10 \times 10 \times 10 = 9000$ எண்கள்

(ஆ) $9 \times 9 \times 8 \times 7 = 4536$ எண்கள்

(இ) $9 \times 8 \times 7 = 504$ எண்கள்

28. மொத்தமுள்ள 500 மாணவர்களில் கணிதம் படித்தவர்கள் 329, புள்ளியியல் படித்தவர்கள் = 295, பொருளியல் படித்தவர்கள் = 186, கணிதமும் பொருளியலும் படித்தவர்கள் = 83, கணிதமும் புள்ளியியலும் படித்தவர்கள் = 217. பொருளியலும் புள்ளியியலும் படித்தவர்கள் = 63. எத்தனை மாணவர்கள் 3 பாடங்களையும் படித்தவர்கள்? மற்ற சேர்வைகளைப் (combinations) படித்தவர்கள் எத்தனை பேர்?

பதில் : இதற்குப் பதில் கண்டுபிடிக்கக் கீழ்வரும் சூத்திரத்தைப் பயன்படுத்தலாம்.



(A+B+C) = (A) + (B)+(C) - (AB) - (BC) - (AC)+(ABC)

(ABC) = 53.

29. இந்தியாவில் ஒரு நகரில் ஓர் ஆண்டில் 200 நாட்களில் நடந்த பெண் சிசுக்கொலை விபரங்கள் கீழே தரப்படுகின்றன. அதற்கு பாய்சான் (POISSON) பரவலைக் காணவும்.

பெண்சிசுக்கொலை எண்ணிக்கை	0	1	2	3	4
நாட்கள் (அலைவெண்கள்)	109	65	22	3	1

பதில் : $P(X) = \dfrac{e^{-0.61}(0.61)}{X!}$

எதிர்பார்க்கப்படும் அலைவெண்கள் முறையே 108.7, 66.3, 20.2, 4.1, 0.7 ஆகும்.

30. ஒரு முழுமையில் 2, 3, 6, 8, 11 ஆகிய ஐந்து எண்கள் உள்ளன. மீண்டும் சேர்க்கும் முறையைப் (with replacement) பயன்படுத்தி இரண்டு எண்கள் கொண்ட மாதிரியாக (அ) எத்தனை மாதிரிகள் எடுக்கலாம்? (ஆ)அவை யாவை (இ) அந்த மாதிரிகளின் கூட்டுச் சராசரிகள் யாவை? (ஈ) முழுமையின் கூட்டுச் சராசரி யாது? (உ) முழுமையின் மாறுபாடு (variance) யாது (ஊ)மாதிரிகளின் கூட்டுச் சராசரிகளின் மாறுபாடு யாது? (எ)மாதிரிகளின் கூட்டுச் சராசரிகளின் கூட்டுச் சராசரி யாது? (ஏ) 'உ' வுக்கும் 'ஊ' வுக்கும் உள்ள தொடர்பு யாது?

பதில் : (அ) 25 மாதிரிகள் (5^2) எடுக்கலாம்.

(ஆ) (2,2) (3, 2) (6,2) (8,2) (11,2)

(2,3) (3, 3) (6,3) (8,3) (11,3)

(2,6) (3, 6) (6,6) (8,6) (11,6)

(2,8) (3, 8) (6,8) (8,8) (11,8)

(2,11) (3, 11) (6,11) (8,11) (11,11)

(இ) (2.0) (2.5) (4.0) (5.0) (6.5)

(2.5) (3.0) (4.5) (5.5) (7.0)

(4.0) (4.5) (6.0) (7.0) (8.5)

(5.0) (5.5) (7.0) (8.0) (9.5)

(6.5) (7.0) (8.5) (9.5) (11.0)

(ஈ) $\frac{2+3+6+8+11}{5} = 6.0$

(உ) $\sigma^2 = \frac{16+9+0+4+25}{5} = 10.8$

(ஊ) $\sigma_{\bar{x}} = \frac{135}{25} = 5.40$

(எ) $\mu_{\bar{x}} = \frac{150}{25} = 6.0$ (இது முழுமையின் சராசரிக்குச் சமமாக இருப்பதைப் பார்க்கலாம்).

(ஏ) $\sigma^2_{\bar{x}} = \frac{\sigma^2}{n}$

$5.400 = \frac{10.8}{2} \quad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$

$\sqrt{5.40} = \frac{\sqrt{10.8}}{\sqrt{2}} = \frac{3.29}{1.414} = 2.32$

2.32 = 2.32

31. *பாரபட்சமற்ற திறன் படைத்த* (unbiased and efficient) *மதிப்பீடுகளுக்கு உதாரணங்கள் யாவை?*

பதில் : *மாதிரியின் கூட்டுச் சராசரியும்* ($\bar{X}$) *மாதிரியின் மாற்றப்பட்ட மாறுபாடும்* (modified sample variance) $\hat{s}^2 = \frac{n}{n-1} s^2$ ஆகும்.

32. பாரபட்சமற்ற மற்றும் திறனற்ற (unbiased and inefficient) மதிப்பீடுகளுக்கு உதாரணங்கள் யாவை?

பதில் : மாதிரியின் இடைநிலையும் (sample median) மாதிரி புள்ளியான $\frac{1}{2}(Q_1 + Q_3)$யும் ஆகும்.

33. பாரபட்சமுள்ள மற்றும் திறனற்ற (biased and inefficient) மதிப்பீடுகளுக்கு உதாரணங்கள் தருக.

பதில் : மாதிரியின் திட்டவிலக்கம் (s), மாதிரியின் மாற்றப்பட்ட திட்ட விலக்கம் ($\hat{s}$), சராசரி விலக்கம் (mean deviation) கால்மானங்களின் வீச்சின் பாதி (semi-interquartile range) ஆகியவை.

34. 100 நல்ல நாணயங்களை 100 தடவைகள் குலுக்கிப் போட்டால் எத்தனை தடவைகள் தலை 40க்கு மேலாகவும் 60க்குள்ளாகவும் வரலாம்?

பதில் : இதற்கு, முதலில் தலை 40க்கு மேலாகவும் 60க்குள்ளாகவும் வருவதற்கான நிகழ்தகவைக் கண்டுபிடித்து அதனை 100ஆல் பெருக்க வேண்டும். அதற்கான நிகழ்தகவை ஈருறுப்புப் பரவல் மூலம் பெற

$$100_{c_{40}} \left(\frac{1}{2}\right)^{40}\left(\frac{1}{2}\right)^{60} + 100_{c_{41}} \left(\frac{1}{2}\right)^{41} \left(\frac{1}{2}\right)^{59} + \ldots$$

$+ 100_{c_{60}} \left(\frac{1}{2}\right)^{60} \left(\frac{1}{2}\right)^{40}$ எனக் கண்டுபிடிக்க வேண்டும். ஆனால், இது சிரமமாக இருக்கும். எனவே, வேறு வழியினைத் தேட வேண்டும். இங்கு $np = 100\left(\frac{1}{2}\right)$

மேலும் $nq = 100\left(\frac{1}{2}\right)$ என 5க்கும் மேலாக இருப்பதால் இங்கு ஈருறுப்புப் பரவலைத் தோராயப்படுத்தி இயல்நிலைப் பரவலாக ஆக்கலாம். இயல்நிலைப் பரவல் தொடர் மாறிகளுக்கானதாகையால் (Continuous

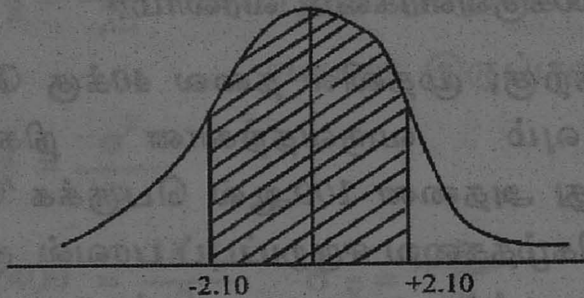
variable), 40 முதல் 60யையும் தொடர் மாறியாக மாற்ற வேண்டும். அதற்காக 39.5 முதல் 60.5 வரை என எடுத்துக் கொள்ளலாம். இந்த எண்களை நிலைப்படுத்தப்பட்ட அலகுகளாக (standardized units) மாற்ற கூட்டுச் சராசரியையும் திட்ட விலக்கத்தையும் காணவேண்டும்.

கூட்டுச் சராசரி $= np = 100\left(\frac{1}{2}\right) = 50$

திட்டவிலக்கம் $= \sqrt{npq} = \sqrt{100\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)} = 5$

39.5ன் நிலைப்படுத்தப்பட்ட அலகு $= \frac{39.5 - 50}{5} = -2.10$

60.5ன் நிலைப்படுத்தப்பட்ட அலகு $= \frac{60.5 - 50}{5} = 2.10$



தேவையான நிகழ்தகவினை வரைபடம் மூலம் காட்டலாம். அது நிழலாக்கப்பட்ட பகுதியாகும். இதனை புள்ளியியல் பட்டியலில் பார்த்தால் இந்தப் பரப்பளவு 0.48 + 0.48 = 0.96 ஆக உள்ளது. எனவே, 39.5 மேலாகவும் 60.5க்குள்ளாகவும் தலைகள் வருவதற்கான நிகழ்தகவு 0.96 ஆகும். 100 தடவைகள் நாணயங்களைக் குலுக்கிப் போடுவதால், 96 தடவைகள் (0.96 × 100) 40க்கு மேல் 60க்குக் கீழ் தலைகள் வரலாம்.

35. டெட்ராகோரிக் (TETRACHORIC) ஒட்டுறவு என்றால் என்ன?

பதில் : ஓர் 2x2 நேர்வுப்பட்டியலில் கொடுக்கப்பட்டுள்ள பண்புகளுக்கிடையேயான உறவினை χ^2 மூலம் அளந்து அதன் மூலம் கண்டுபிடிக்கப்படும் ஒட்டுறவுக்கு டெட்ராகோரிக் ஒட்டுறவு என்று பெயர். அதற்கான சூத்திரம்.

$$r = \sqrt{\frac{\chi^2}{n\,(k-1)}}$$

36. *(2, -3)* மற்றும் *(4,5)* ஆகிய புள்ளிகளைத் தொட்டுச் செல்லும் நேர்கோட்டிற்கான சமன்பாட்டினைத் தருவிக்கவும்.

பதில் : $X_1 = 2 \quad Y_1 = -3$

$X_2 = 4 \quad Y_2 = 5$

$$(m) = \text{சாய்வு} = \frac{Y_2 - Y_1}{X_2 - X_1} = \frac{5 - (-3)}{4 - 2} = \frac{8}{2} = 4$$

$$Y - Y_1 = m\,(X - X_1)$$

$$Y - (-3) = 4\,(X-2)$$

$$Y + 3 = 4X - 8$$

$$Y = 4X - 8 - 3$$

$$Y = 4X - 11$$

இதில் Y = சார்பு மாறி; X = சாரா மாறி; *4* = சாய்வு, *11* = வெட்டுத்துண்டு (intercept).

37. (4,2) புள்ளியைத் தொட்டும் 2X + 3Y = 6 என்ற நேர்கோட்டுக்கு இணையாகவும் உள்ள நேர்கோட்டைக் காணவும்.

பதில் : இரண்டு இணையான நேர்கோடுகளின் சாய்வுகளும் (slopes) சமமாக இருக்க வேண்டும்.

2X + 3Y = 6 எனில் அதன் சாய்வு $-\frac{2}{3}$ ஆகும்.

$$\left[\begin{array}{l} 3Y = 6 - 2X \\ Y = 2 - \frac{2}{3}X \end{array}\right]$$

எனவே, $Y - 2 = -\frac{2}{3}(X - 4)$

$$Y = -\frac{2}{3}X + \frac{8}{3} + 2 \qquad \left[\begin{array}{l} \frac{8}{3} + \frac{2}{1} = \frac{8+6}{3} \\ \qquad\quad = \frac{14}{3} \end{array}\right]$$

$$Y = -\frac{2}{3}X + \frac{14}{3}$$

$$3Y = -2X + 14$$

3Y + 2X = 14 அல்லது

$$2X + 3Y = 14$$

38. கீழ்க்காணும் விபரங்களைப் பயன்படுத்தி, மீச்சிறு வர்க்கப் பரவளைவு (Least Square Parabola) காணவும்.

$\Sigma X = 0$ $\quad \Sigma X^2 = 110$ $\quad \Sigma X^4 = 1958$ $\quad \Sigma X^2Y = 9209.0$

$\Sigma Y = 886.8$ $\quad \Sigma X^3 = 0$ $\quad \Sigma XY = 1429.8$ $\quad n = 11$

பதில் :

$$\Sigma Y = aN + b_1\Sigma X + b_2\Sigma X^2$$

$$\Sigma XY = a\Sigma X + b_1\Sigma X^2 + b_2\Sigma X^3$$

$\Sigma X^2Y = a\Sigma X^2 + b_1\Sigma X^3 + b_2\Sigma X^4$

$886.8 = 11a + 110b_2$

$1429.8 = 110b_1$

$9209.0 = 110a + 1958b_2$

இவற்றிலிருந்து, $Y = 76.64 + 13.00X + 0.3974X^2$

39. 18 கூறுகளைக் கொண்ட ஒரு மாதிரியில் உள்ள விபரங்களுக்கு இடையேயான ஒட்டுறவுக் கெழு 0.32. இது தொடர்பான முழுமையின் ஒட்டுறவுக்கெழு (Population Correlation Coefficient), பூஜ்யத்தைவிட்டு விலகி உள்ளதென 0.05 அளவான புள்ளியியல் முக்கியத்துடன் சொல்ல முடியுமா?

பதில் : இல்லெனும் எடுகோள் $H_0 : \rho = 0$

மாற்று எடுகோள் : $H_1 : \rho > 0$

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0.32\sqrt{18-2}}{\sqrt{1-(0.32)^2}} = 1.35$$

இதற்குப் பொருத்தமான tயின் அட்டவணை மதிப்பு 1.75 ஆக இருப்பதால் இல்லெனும் எடுகோளை ஒத்துக் கொள்ள வேண்டியுள்ளது. எனவே, முழுமையின் ஒட்டுறவுக்கெழு பூஜ்யமாக இருக்கலாம்.

40. ஒரு வகுப்பில் உள்ள 100 மாணவர்களின் மதிப்பெண்கள் இயல்நிலைப் பரவலாக உள்ளது. அந்த 100 மாணவர்கள் பெற்ற மதிப்பெண்களின் கூட்டுச் சராசரி 62.73 எனில், இடைநிலை யாது? முகடு யாது?

பதில் : ஓர் இயல்நிலைப் பரவலின் கூட்டுச் சராசரி, இடைநிலை, முகடு ஆகியவை சமமாக இருக்கும் என்பதால், இடைநிலையும் முகடும் 62.73 தான்.

N	0	1	2	3	4	5	6	7	8	9
10	0000	0043	0086	0128	0170	0212	0253	0294	0334	0374
11	0414	0453	0492	0531	0569	0607	0645	0682	0719	0755
12	0792	0828	0864	0899	0934	0969	1004	1038	1072	1106
13	1139	1173	1206	1239	1271	1303	1335	1367	1399	1430
14	1461	1492	1523	1553	1584	1614	1644	1673	1703	1732
15	1761	1790	1818	1847	1875	1903	1931	1959	1987	2014
16	2041	2068	2095	2122	2148	2175	2201	2227	2253	2279
17	2304	2330	2355	2380	2405	2430	2455	2480	2504	2529
18	2553	2577	2601	2625	2648	2672	2695	2718	2742	2765
19	2788	2810	2833	2856	2878	2900	2923	2945	2967	2989
20	3010	3032	3054	3075	3096	3118	3139	3160	3181	3201
21	3222	3243	3263	3284	3304	3324	3345	3365	3385	3404
22	3424	3444	3464	3483	3502	3522	3541	3560	3579	3598
23	3617	3636	3655	3674	3692	3711	3729	3747	3766	3784
24	3802	3820	3838	3856	3874	3892	3909	3927	3945	3962
25	3979	3997	4014	4031	4048	4065	4082	4099	4116	4133
26	4150	4166	4183	4200	4216	4232	4249	4265	4281	4298
27	4314	4330	4346	4362	4378	4393	4409	4425	4440	4456
28	4472	4487	4502	4518	4533	4548	4564	4579	4594	4609
29	4624	4639	4654	4669	4683	4698	4713	4728	4742	4757
30	4771	4786	4800	4814	4829	4843	4857	4871	4886	4900
31	4914	4928	4942	4955	4969	4983	4997	5011	5024	5038
32	5051	5065	5079	5092	5105	5119	5132	5145	5159	5172
33	5185	5198	5211	5224	5237	5250	5263	5276	5289	5302
34	5315	5328	5340	5353	5366	5378	5391	5403	5416	5428
35	5441	5453	5465	5478	5490	5502	5514	5527	5539	5551
36	5563	5575	5587	5599	5611	5623	5635	5647	5658	5670
37	5682	5694	5705	5717	5729	5740	5752	5763	5775	5786
38	5798	5809	5821	5832	5843	5855	5866	5877	5888	5899
39	5911	5922	5933	5944	5955	5966	5977	5988	5999	6010
40	6021	6031	6042	6053	6064	6075	6085	6096	6107	6117
41	6128	6138	6149	6160	6170	6180	6191	6201	6212	6222
42	6232	6243	6253	6263	6274	6284	6294	6304	6314	6325
43	6335	6345	6355	6365	6375	6385	6395	6405	6415	6425
44	6435	6444	6454	6464	6474	6484	6493	6503	6513	6522
45	6532	6542	6551	6561	6571	6580	6590	6599	6609	6618
46	6628	6637	6646	6656	6665	6675	6684	6693	6702	6712
47	6721	6730	6739	6749	6758	6767	6776	6785	6794	6803
48	6812	6821	6830	6839	6848	6857	6866	6875	6884	6893
49	6902	6911	6920	6928	6937	6946	6955	6964	6972	6981
50	6990	6998	7007	7016	7024	7033	7042	7050	7059	7067
51	7076	7084	7093	7101	7110	7118	7126	7135	7143	7152
52	7160	7168	7177	7185	7193	7202	7210	7218	7226	7235
53	7243	7251	7259	7267	7275	7284	7292	7300	7308	7316
54	7324	7332	7340	7348	7356	7364	7372	7380	7388	7396

புள்ளியியல் அட்டவணை - 1ஆ (Logarithms)

N	0	1	2	3	4	5	6	7	8	9
55	7404	7412	7419	7427	7435	7443	7451	7459	7466	7474
56	7482	7490	7497	7505	7513	7520	7528	7536	7543	7551
57	7559	7566	7574	7582	7589	7597	7604	7612	7619	7627
58	7634	7642	7649	7657	7664	7672	7679	7686	7694	7701
59	7709	7716	7723	7731	7738	7745	7752	7760	7767	7774
60	7782	7789	7796	7803	7810	7818	7825	7832	7839	7846
61	7853	7860	7868	7875	7882	7889	7896	7903	7910	7917
62	7924	7931	7938	7945	7952	7959	7966	7973	7980	7987
63	7993	8000	8007	8014	8021	8028	8035	8041	8048	8055
64	8062	8069	8075	8082	8089	8096	8102	8109	8116	8122
65	8129	8136	8142	8149	8156	8162	8169	8176	8182	8189
66	8195	8202	8209	8215	8222	8228	8235	8241	8248	8254
67	8261	8267	8274	8280	8287	8293	8299	8306	8312	8319
68	8325	8331	8338	8344	8351	8357	8363	8370	8376	8382
69	8388	8395	8401	8407	8414	8420	8426	8432	8439	8445
70	8451	8457	8463	8470	8476	8482	8488	8494	8500	8506
71	8513	8519	8525	8531	8537	8543	8549	8555	8561	8567
72	8573	8579	8585	8591	8597	8603	8609	8615	8621	8627
73	8633	8639	8645	8651	8657	8663	8669	8675	8681	8686
74	8692	8698	8704	8710	8716	8722	8727	8733	8739	8745
75	8751	8756	8762	8768	8774	8779	8785	8791	8797	8802
76	8808	8814	8820	8825	8831	8837	8842	8848	8854	8859
77	8865	8871	8876	8882	8887	8893	8899	8904	8910	8915
78	8921	8927	8932	8938	8943	8949	8954	8960	8965	8971
79	8976	8982	8987	8993	8998	9004	9009	9015	9020	9025
80	9031	9036	9042	9047	9053	9058	9063	9069	9074	9079
81	9085	9090	9096	9101	9106	9112	9117	9122	9128	9133
82	9138	9143	9149	9154	9159	9165	9170	9175	9180	9186
83	9191	9196	9201	9206	9212	9217	9222	9227	9232	9238
84	9243	9248	9253	9258	9263	9269	9274	9279	9284	9289
85	9294	9299	9304	9309	9315	9320	9325	9330	9335	9340
86	9345	9350	9355	9360	9365	9370	9375	9380	9385	9390
87	9395	9400	9405	9410	9415	9420	9425	9430	9435	9440
88	9445	9450	9455	9460	9465	9469	9474	9479	9484	9489
89	9494	9499	9504	9509	9513	9518	9523	9528	9533	9538
90	9542	9547	9552	9557	9562	9566	9571	9576	9581	9586
91	9590	9595	9600	9605	9609	9614	9619	9624	9628	9633
92	9638	9643	9647	9652	9657	9661	9666	9671	9675	9680
93	9685	9689	9694	9699	9703	9708	9713	9717	9722	9727
94	9731	9736	9741	9745	9750	9754	9759	9763	9768	9773
95	9777	9782	9786	9791	9795	9800	9805	9809	9814	9818
96	9823	9827	9832	9836	9841	9845	9850	9854	9859	9863
97	9868	9872	9877	9881	9886	9890	9894	9899	9903	9908
98	9912	9917	9921	9926	9930	9934	9939	9943	9948	9952
99	9956	9961	9965	9969	9974	9978	9983	9987	9991	9996

											Mean Differences								
	0	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9
·00	1000	1002	1005	1007	1009	1012	1014	1016	1019	1021	0	0	1	1	1	1	2	2	2
·01	1023	1026	1028	1030	1033	1035	1038	1040	1042	1045	0	0	1	1	1	1	2	2	2
·02	1047	1050	1052	1054	1057	1059	1062	1064	1067	1069	0	0	1	1	1	1	2	2	2
·03	1072	1074	1076	1079	1081	1084	1086	1089	1091	1094	0	0	1	1	1	1	2	2	2
·04	1096	1099	1102	1104	1107	1109	1112	1114	1117	1119	0	1	1	1	1	2	2	2	2
·05	1122	1125	1127	1130	1132	1135	1138	1140	1143	1146	0	1	1	1	1	2	2	2	2
·06	1148	1151	1153	1156	1159	1161	1164	1167	1169	1172	0	1	1	1	1	2	2	2	2
·07	1175	1178	1180	1183	1186	1189	1191	1194	1197	1199	0	1	1	1	1	2	2	2	2
·08	1202	1205	1208	1211	1213	1216	1219	1222	1225	1227	0	1	1	1	1	2	2	2	3
·09	1230	1233	1236	1239	1242	1245	1247	1250	1253	1256	0	1	1	1	1	2	2	2	3
·10	1259	1262	1265	1268	1271	1274	1276	1279	1282	1285	0	1	1	1	1	2	2	2	3
·11	1288	1291	1294	1297	1300	1303	1306	1309	1312	1315	0	1	1	1	2	2	2	2	3
·12	1318	1321	1324	1327	1330	1334	1337	1340	1343	1346	0	1	1	1	2	2	2	2	3
·13	1349	1352	1355	1358	1361	1365	1368	1371	1374	1377	0	1	1	1	2	2	2	3	3
·14	1380	1384	1387	1390	1393	1396	1400	1403	1406	1409	0	1	1	1	2	2	2	3	3
·15	1413	1416	1419	1422	1426	1429	1432	1435	1439	1442	0	1	1	1	2	2	2	3	3
·16	1445	1449	1452	1455	1459	1462	1466	1469	1472	1476	0	1	1	1	2	2	2	3	3
·17	1479	1483	1486	1489	1493	1496	1500	1503	1507	1510	0	1	1	1	2	2	2	3	3
·18	1514	1517	1521	1524	1528	1531	1535	1538	1542	1545	0	1	1	1	2	2	2	3	3
·19	1549	1552	1556	1560	1563	1567	1570	1574	1578	1581	0	1	1	1	2	2	3	3	3
·20	1585	1589	1592	1596	1600	1603	1607	1611	1614	1618	0	1	1	1	2	2	3	3	3
·21	1622	1626	1629	1633	1637	1641	1644	1648	1652	1656	0	1	1	2	2	2	3	3	3
·22	1660	1663	1667	1671	1675	1679	1683	1687	1690	1694	0	1	1	2	2	2	3	3	3
·23	1698	1702	1706	1710	1714	1718	1722	1726	1730	1734	0	1	1	2	2	2	3	3	4
·24	1738	1742	1746	1750	1754	1758	1762	1766	1770	1774	0	1	1	2	2	2	3	3	4
·25	1778	1782	1786	1791	1795	1799	1803	1807	1811	1816	0	1	1	2	2	2	3	3	4
·26	1820	1824	1828	1832	1837	1841	1845	1849	1854	1858	0	1	1	2	2	3	3	3	4
·27	1862	1866	1871	1875	1879	1884	1888	1892	1897	1901	0	1	1	2	2	3	3	3	4
·28	1905	1910	1914	1919	1923	1928	1932	1936	1941	1945	0	1	1	2	2	3	3	4	4
·29	1950	1954	1959	1963	1968	1972	1977	1982	1986	1991	0	1	1	2	2	3	3	4	4
·30	1995	2000	2004	2009	2014	2018	2023	2028	2032	2037	0	1	1	2	2	3	3	4	4
·31	2042	2046	2051	2056	2061	2065	2070	2075	2080	2084	0	1	1	2	2	3	3	4	4
·32	2089	2094	2099	2104	2109	2113	2118	2123	2128	2133	0	1	1	2	2	3	3	4	4
·33	2138	2143	2148	2153	2158	2163	2168	2173	2178	2183	0	1	1	2	2	3	3	4	4
·34	2188	2193	2198	2203	2208	2213	2218	2223	2228	2234	1	1	2	2	3	3	4	4	5
·35	2239	2244	2249	2254	2259	2265	2270	2275	2280	2286	1	1	2	2	3	3	4	4	5
·36	2291	2296	2301	2307	2312	2317	2323	2328	2333	2339	1	1	2	2	3	3	4	4	5
·37	2344	2350	2355	2360	2366	2371	2377	2382	2388	2393	1	1	2	2	3	3	4	4	5
·38	2399	2404	2410	2415	2421	2427	2432	2438	2443	2449	1	1	2	2	3	3	4	4	5
·39	2455	2460	2466	2472	2477	2483	2489	2495	2500	2506	1	1	2	2	3	3	4	5	5
·40	2512	2518	2523	2529	2535	2541	2547	2553	2559	2564	1	1	2	2	3	4	4	5	5
·41	2570	2576	2582	2588	2594	2600	2606	2612	2618	2624	1	1	2	2	3	4	4	5	5
·42	2630	2636	2642	2649	2655	2661	2667	2673	2679	2685	1	1	2	2	3	4	4	5	6
·43	2692	2698	2704	2710	2716	2723	2729	2735	2742	2748	1	1	2	3	3	4	4	5	6
·44	2754	2761	2767	2773	2780	2786	2793	2799	2805	2812	1	1	2	3	3	4	4	5	6
·45	2818	2825	2831	2838	2844	2851	2858	2864	2871	2877	1	1	2	3	3	4	5	5	6
·46	2884	2891	2897	2904	2911	2917	2924	2931	2938	2944	1	1	2	3	3	4	5	5	6
·47	2951	2958	2965	2972	2979	2985	2992	2999	3006	3013	1	1	2	3	3	4	5	5	6
·48	3020	3027	3034	3041	3048	3055	3062	3069	3076	3083	1	1	2	3	4	4	5	6	6
·49	3090	3097	3105	3112	3119	3126	3133	3141	3148	3155	1	1	2	3	4	4	5	6	7

											Mean Differences								
	0	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9
50	3162	3170	3177	3184	3192	3199	3206	3214	3221	3228	1	1	2	3	4	4	5	6	7
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

Squares, Square Roots, and Reciprocals

n	n^2	$\sqrt{n}$	$\sqrt{10n}$	$1/n$	n	n^2	$\sqrt{n}$	$\sqrt{10n}$	$1/n$
1	1	1.000	3.162	1.00000	51	2601	7.141	22.583	.01961
2	4	1.414	4.472	.50000	52	2704	7.211	22.804	.01923
3	9	1.732	5.477	.33333	53	2809	7.280	23.022	.01887
4	16	2.000	6.325	.25000	54	2916	7.348	23.238	.01852
5	25	2.236	7.071	.20000	55	3025	7.416	23.452	.01818
6	36	2.449	7.746	.16667	56	3136	7.483	23.664	.01786
7	49	2.646	8.367	.14286	57	[illegible]	[illegible]	[illegible]	[illegible]
8	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
9	81	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
10	100	3.162	10.000	.10000	60	3600	7.746	24.495	.01667
11	121	3.317	10.488	.09091	61	3721	7.810	24.698	.01639
12	144	3.464	10.954	.08333	62	3844	7.874	24.900	.01613
13	169	3.606	11.402	.07692	63	3969	7.937	25.100	.01587
14	196	3.742	11.832	.07143	64	4096	8.000	25.298	.01562
15	225	3.873	12.247	.06667	65	4225	8.062	25.495	.01538
16	256	4.000	12.649	.06250	66	4356	8.124	25.690	.01515
17	289	4.123	13.038	.05882	67	4489	8.185	25.884	.01493
18	324	4.243	13.416	.05556	68	4624	8.246	26.077	.01471
19	361	4.359	13.784	.05263	69	4761	8.307	26.268	.01449
20	400	4.472	14.142	.05000	70	4900	8.367	26.458	.01429
21	441	4.583	14.491	.04762	71	5041	8.426	26.646	.01408
22	484	4.690	14.832	.04545	72	5184	8.485	26.833	.01389
23	529	4.796	15.166	.04348	73	5329	8.544	27.019	.01370
24	576	4.899	15.492	.04167	74	5476	8.602	27.203	.01351
25	625	5.000	15.811	.04000	75	5625	8.660	27.386	.01333
26	676	5.099	16.125	.03846	76	5776	8.718	27.568	.01316
27	729	5.196	16.432	.03704	77	5929	8.775	27.749	.01299
28	784	5.292	16.733	.03571	78	6084	8.832	27.928	.01282
29	841	5.385	17.029	.03448	79	6241	8.888	28.107	.01266
30	900	5.477	17.321	.03333	80	6400	8.944	28.284	.01250
31	961	5.568	17.607	.03226	81	6561	9.000	28.460	.01235
32	1024	5.657	17.889	.03125	82	6724	9.055	28.636	.01220
33	1089	5.745	18.166	.03030	83	6889	9.110	28.810	.01205
34	1156	5.831	18.439	.02941	84	7056	9.165	28.983	.01190
35	1225	5.916	18.708	.02857	85	7225	9.220	29.155	.01176
36	1296	6.000	18.974	.02778	86	7396	9.274	29.326	.01163
37	1369	6.083	19.235	.02703	87	7569	9.327	29.496	.01149
38	1444	6.164	19.494	.02632	88	7744	9.381	29.665	.01136
39	1521	6.245	19.748	.02564	89	7921	9.434	29.833	.01124
40	1600	6.325	20.000	.02500	90	8100	9.487	30.000	.01111
41	1681	6.403	20.248	.02439	91	8281	9.539	30.166	.01099
42	1764	6.481	20.494	.02381	92	8464	9.592	30.332	.01087
43	1849	6.557	20.736	.02326	93	8649	9.644	30.496	.01075
44	1936	6.633	20.976	.02273	94	8836	9.695	30.659	.01064
45	2025	6.708	21.213	.02222	95	9025	9.747	30.822	.01053
46	2116	6.782	21.448	.02174	96	9216	9.798	30.984	.01042
47	2209	6.856	21.679	.02128	97	9409	9.849	31.145	.01031
48	2304	6.928	21.909	.02083	98	9604	9.899	31.305	.01020
49	2401	7.000	22.136	.02041	99	9801	9.950	31.464	.01010
50	2500	7.071	22.361	.02000	100	10000	10.000	31.623	.01000

புள்ளியியல் அட்டவணை - 4 (Random Numbers)

51772	74640	42331	29044	46621	62898	93582	04186	19640	87056
24033	23491	83587	06568	21960	21387	76105	10863	97453	90581
45939	60173	52078	25424	11645	55870	56974	37428	93507	94271
30586	02133	75797	45406	31041	86707	12973	17169	88116	42187
03585	79353	81938	82322	96799	85659	36081	50884	14070	74950
64937	03355	95863	20790	65304	55189	00745	65253	11822	15804
15630	64759	51135	98527	62586	41889	25439	88036	24034	67283
09448	56301	57683	30277	94628	85418	68829	06652	41982	49159
21631	91157	77331	60710	52290	16885	48658	71590	16159	14676
91097	17480	29414	06829	87843	28195	27279	47152	35683	47280
50532	25496	95652	42457	73547	76552	50020	24819	52984	76168
07136	40876	79971	54195	25708	51817	36732	72484	94923	75936
27989	64728	10744	08396	56242	90985	28868	99431	50995	20507
85184	73949	36601	46253	00477	25234	09908	36574	72139	70185
54398	21154	97810	36764	32869	11785	55261	59009	38714	38723
65544	34371	09591	07839	58892	92843	72828	91341	84821	63886
08263	65952	85762	64286	39288	18776	84303	99247	46149	03229
39817	67906	48236	16057	81812	15815	63700	85915	19219	45943
62257	04077	79443	95203	02479	30763	92486	54083	23631	05825
53298	90276	62545	21944	16530	03878	07516	95715	02526	83537

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990

Also, for $z = 4.0$, 5.0, and 6.0, the areas are 0.49997, 0.4999997, and 0.499999999.



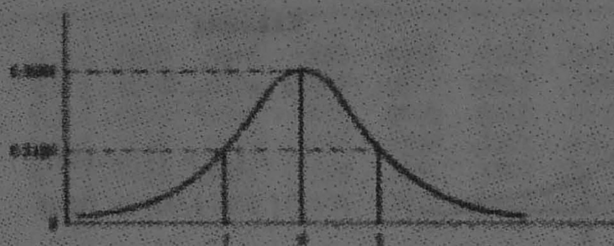
Example

$$z = \frac{X - \mu}{\sigma}$$

$P[z > 1] = .1587$

$P[z > 1.96] = .0250$

Normal Deviate z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010



Example

$$z = \frac{x - \mu}{\sigma}$$

$z = 0$: ordinate $= 0.3989$

$z = 1$: ordinate $= 0.2420$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.3989	.3989	.3989	.3988	.3986	.3984	.3982	.3980	.3977	.3973
.1	.3970	.3965	.3961	.3956	.3951	.3945	.3939	.3932	.3925	.3918
.2	.3910	.3902	.3894	.3885	.3876	.3867	.3857	.3847	.3836	.3825
.3	.3814	.3802	.3790	.3778	.3765	.3752	.3739	.3725	.3712	.3697
.4	.3683	.3668	.3653	.3637	.3621	.3605	.3589	.3572	.3555	.3538
.5	.3521	.3503	.3485	.3467	.3448	.3429	.3410	.3391	.3372	.3352
.6	.3332	.3312	.3292	.3271	.3251	.3230	.3209	.3187	.3166	.3144
.7	.3123	.3101	.3079	.3056	.3034	.3011	.2989	.2966	.2943	.2920
.8	.2897	.2874	.2850	.2827	.2803	.2780	.2756	.2732	.2709	.2685
.9	.2661	.2637	.2613	.2589	.2565	.2541	.2516	.2492	.2468	.2444
1.0	.2420	.2396	.2371	.2347	.2323	.2299	.2275	.2251	.2227	.2203
1.1	.2179	.2155	.2131	.2107	.2083	.2059	.2036	.2012	.1989	.1965
1.2	.1942	.1919	.1895	.1872	.1849	.1826	.1804	.1781	.1758	.1736
1.3	.1714	.1691	.1669	.1647	.1626	.1604	.1582	.1561	.1539	.1518
1.4	.1497	.1476	.1456	.1435	.1415	.1394	.1374	.1354	.1334	.1315
1.5	.1295	.1276	.1257	.1238	.1219	.1200	.1182	.1163	.1145	.1127
1.6	.1109	.1092	.1074	.1057	.1040	.1023	.1006	.0989	.0973	.0957
1.7	.0940	.0925	.0909	.0893	.0878	.0863	.0848	.0833	.0818	.0804
1.8	.0790	.0775	.0761	.0748	.0734	.0721	.0707	.0694	.0681	.0669
1.9	.0656	.0644	.0632	.0620	.0608	.0596	.0584	.0573	.0562	.0551
2.0	.0540	.0529	.0519	.0508	.0498	.0488	.0478	.0468	.0459	.0449
2.1	.0440	.0431	.0422	.0413	.0404	.0396	.0387	.0379	.0371	.0363
2.2	.0355	.0347	.0339	.0332	.0325	.0317	.0310	.0303	.0297	.0290
2.3	.0283	.0277	.0270	.0264	.0258	.0252	.0246	.0241	.0235	.0229
2.4	.0224	.0219	.0213	.0208	.0203	.0198	.0194	.0189	.0184	.0180
2.5	.0175	.0171	.0167	.0163	.0158	.0154	.0151	.0147	.0143	.0139
2.6	.0136	.0132	.0129	.0126	.0122	.0119	.0116	.0113	.0110	.0107
2.7	.0104	.0101	.0099	.0096	.0093	.0091	.0088	.0086	.0084	.0081
2.8	.0079	.0077	.0075	.0073	.0071	.0069	.0067	.0065	.0063	.0061
2.9	.0060	.0058	.0056	.0055	.0053	.0051	.0050	.0048	.0047	.0046
3.0	.0044	.0043	.0042	.0040	.0039	.0038	.0037	.0036	.0035	.0034
3.1	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026	.0025	.0025
3.2	.0024	.0023	.0022	.0022	.0021	.0020	.0020	.0019	.0018	.0018
3.3	.0017	.0017	.0016	.0016	.0015	.0015	.0014	.0014	.0013	.0013
3.4	.0012	.0012	.0012	.0011	.0011	.0010	.0010	.0010	.0009	.0009
3.5	.0009	.0008	.0008	.0008	.0008	.0007	.0007	.0007	.0007	.0006
3.6	.0006	.0006	.0006	.0005	.0005	.0005	.0005	.0005	.0005	.0004
3.7	.0004	.0004	.0004	.0004	.0004	.0004	.0003	.0003	.0003	.0003
3.8	.0003	.0003	.0003	.0003	.0003	.0002	.0002	.0002	.0002	.0002
3.9	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0001	.0001

Example

For ϕ

of ...

p



n	.25	.20	.15	.10	.05	.025	.01		
1	1.000	1.376	1.963	3.078	6.314	12.706	31.821	[illegible]	[illegible]
2	.816	1.061	1.386	1.886	2.920	4.303	6.965	[illegible]	[illegible]
3	.765	.978	1.250	1.638	2.353	3.182	4.541	[illegible]	[illegible]
4	.741	.941	1.190	1.533	2.132	2.776	3.747	4.604	[illegible]
5	.727	.920	1.156	1.476	2.015	2.571	3.365	4.032	6.859
6	.718	.906	1.134	1.440	1.943	2.447	3.143	3.707	5.959
7	.711	.896	1.119	1.415	1.895	2.365	2.998	3.499	5.405
8	.706	.889	1.108	1.397	1.860	2.306	2.896	3.355	5.041
9	.703	.883	1.100	1.383	1.833	2.262	2.821	3.250	4.781
10	.700	.879	1.093	1.372	1.812	2.228	2.764	3.169	4.587
11	.697	.876	1.088	1.363	1.796	2.201	2.718	3.106	4.437
12	.695	.873	1.083	1.356	1.782	2.179	2.681	3.055	4.318
13	.694	.870	1.079	1.350	1.771	2.160	2.650	3.012	4.221
14	.692	.868	1.076	1.345	1.761	2.145	2.624	2.977	4.140
15	.691	.866	1.074	1.341	1.753	2.131	2.602	2.947	4.073
16	.690	.865	1.071	1.337	1.746	2.120	2.583	2.921	4.015
17	.689	.863	1.069	1.333	1.740	2.110	2.567	2.898	3.965
18	.688	.862	1.067	1.330	1.734	2.101	2.552	2.878	3.922
19	.688	.861	1.066	1.328	1.729	2.093	2.539	2.861	3.883
20	.687	.860	1.064	1.325	1.725	2.086	2.528	2.845	3.850
21	.686	.859	1.063	1.323	1.721	2.080	2.518	2.831	3.819
22	.686	.858	1.061	1.321	1.717	2.074	2.508	2.819	3.792
23	.685	.858	1.060	1.319	1.714	2.069	2.500	2.807	3.767
24	.685	.857	1.059	1.318	1.711	2.064	2.492	2.397	3.745
25	.684	.856	1.058	1.316	1.708	2.060	2.485	2.787	3.725
26	.684	.856	1.058	1.315	1.706	2.056	[illegible]	2.779	3.707
27	.684	.855	1.057	1.314	1.703	2.052	2.473	2.771	3.690
28	.683	.855	1.056	1.313	1.701	2.048	2.467	2.763	3.674
29	.683	.854	1.055	1.311	1.699	2.045	2.462	2.756	3.659
30	.683	.854	1.055	1.310	1.697	2.042	2.457	2.750	3.646
40	.681	.851	1.050	1.303	1.684	2.021	2.423	2.704	3.551
60	.679	.848	1.046	1.296	1.671	2.000	2.390	2.660	3.460
120	.677	.845	1.041	1.289	1.658	1.980	2.358	2.617	3.373
∞	.674	.842	1.036	1.282	1.645	1.960	2.326	2.576	3.291

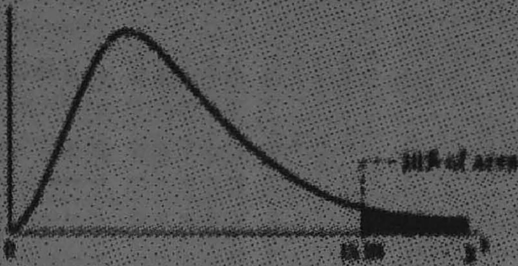
Source: This table is abridged from Table III of Fisher & Yates: *Statistical Tables for Biological, Agricultural and Medical Research* published by Oliver & Boyd Ltd., Edinburgh, and by permission of the authors and publishers.

df \ t_p	.75	.90	.95	.975	.99	.995	.9995
1	1.000	3.078	6.314	12.706	31.821	63.657	636.619
2	.816	1.886	2.920	4.303	6.965	9.925	31.598
3	.765	1.638	2.353	3.182	4.541	5.841	12.941
4	.741	1.533	2.132	2.776	3.747	4.604	8.610
5	.727	1.476	2.015	2.571	3.365	4.032	6.859
6	.718	1.440	1.943	2.447	3.143	3.707	5.959
7	.711	1.415	1.895	2.365	2.998	3.499	5.405
8	.706	1.397	1.860	2.306	2.896	3.355	5.041
9	.703	1.383	1.833	2.262	2.821	3.250	4.781
10	.700	1.372	1.812	2.228	2.764	3.169	4.587
11	.697	1.363	1.796	2.201	2.718	3.106	4.437
12	.695	1.356	1.782	2.179	2.681	3.055	4.318
13	.694	1.350	1.771	2.160	2.650	3.012	4.221
14	.692	1.345	1.761	2.145	2.624	2.977	4.140
15	.691	1.341	1.753	2.131	2.602	2.947	4.073
16	.690	1.337	1.746	2.120	2.583	2.921	4.015
17	.689	1.333	1.740	2.110	2.567	2.898	3.965
18	.688	1.330	1.734	2.101	2.552	2.878	3.922
19	.688	1.328	1.729	2.093	2.539	2.861	3.883
20	.687	1.325	1.725	2.086	2.528	2.845	3.850
21	.686	1.323	1.721	2.080	2.518	2.831	3.819
22	.686	1.321	1.717	2.074	2.508	2.819	3.792
23	.685	1.319	1.714	2.069	2.500	2.807	3.767
24	.685	1.318	1.711	2.064	2.492	2.797	3.745
25	.684	1.316	1.708	2.060	2.485	2.787	3.725
26	.684	1.315	1.706	2.056	2.479	2.779	3.707
27	.684	1.314	1.703	2.052	2.473	2.771	3.690
28	.683	1.313	1.701	2.048	2.467	2.763	3.674
29	.683	1.311	1.699	2.045	2.462	2.756	3.659
30	.683	1.310	1.697	2.042	2.457	2.750	3.646
40	.681	1.303	1.684	2.021	2.423	2.704	3.551
60	.679	1.296	1.671	2.000	2.390	2.660	3.460
120	.677	1.289	1.658	1.980	2.358	2.617	3.373
∞	.674	1.282	1.645	1.960	2.326	2.576	3.291

புள்ளியியல் அட்டவணை - 7 (Percentage Points of the C^2 Distribution)

TABLE 4

Percentage Points of the χ^2 Distribution



Example

For ϕ = 10 degrees of freedom:

$P[\chi^2 > 15.99] = .10$

ϕ \ P	.995	.99	.975	.95	.90	.75	.50	.25	.10	.05	.025	.01	.005	P / ϕ
1	0.0^4393	0.0^3157	0.0^3982	0.0^2393	0.0158	0.102	0.455	1.323	2.71	3.84	5.02	6.63	7.88	1
2	0.0100	0.0201	0.0506	0.103	0.211	0.575	1.386	2.77	4.61	5.99	7.38	9.21	10.60	2
3	0.0717	0.115	0.216	0.352	0.584	1.213	2.37	4.11	6.25	7.81	9.35	11.34	12.84	3
4	0.207	0.297	0.484	0.711	1.064	1.923	3.36	5.39	7.78	9.49	11.14	13.28	14.86	4
5	0.412	0.554	0.831	1.145	1.610	2.67	4.35	6.63	9.24	11.07	12.83	15.09	16.75	5
6	0.676	0.872	1.237	1.635	2.20	3.45	5.35	7.84	10.64	12.59	14.45	16.81	18.55	6
7	0.989	1.239	1.690	2.17	2.83	4.25	6.35	9.04	12.02	14.07	16.01	18.48	20.3	7
8	1.344	1.646	2.18	2.73	3.49	5.07	7.34	10.22	13.36	15.51	17.53	20.1	22.0	8
9	1.735	2.09	2.70	3.33	4.17	5.90	8.34	11.39	14.68	16.92	19.02	21.7	23.6	9
10	2.16	2.56	3.25	3.94	4.87	6.74	9.34	12.55	15.99	18.31	20.5	23.2	25.2	10
11	2.60	3.05	3.82	4.57	5.58	7.58	10.34	13.70	17.28	19.68	21.9	24.7	26.8	11
12	3.07	3.57	4.40	5.23	6.30	8.44	11.34	14.85	18.55	21.0	23.3	26.2	28.3	12
13	3.57	4.11	5.01	5.89	7.04	9.30	12.34	15.98	19.81	22.4	24.7	27.7	29.8	13
14	4.07	4.66	5.63	6.57	7.79	10.17	13.34	17.12	21.1	23.7	26.1	29.1	31.3	14
15	4.60	5.23	6.26	7.26	8.55	11.04	14.34	18.25	22.3	25.0	27.5	30.6	32.8	15
16	5.14	5.81	6.91	7.96	9.31	11.91	15.34	19.37	23.5	26.3	28.8	32.0	34.3	16
17	5.70	6.41	7.56	8.67	10.09	12.79	16.34	20.5	24.8	27.6	30.2	33.4	35.7	17
18	6.26	7.01	8.23	9.39	10.86	13.68	17.34	21.6	26.0	28.9	31.5	34.8	37.2	18
19	6.84	7.63	8.91	10.12	11.65	14.56	18.34	22.7	27.2	30.1	32.9	36.2	38.6	19
20	7.43	8.26	9.59	10.85	12.44	15.45	19.34	23.8	28.4	31.4	34.2	37.6	40.0	20
21	8.03	8.90	10.28	11.59	13.24	16.34	20.3	24.9	29.6	32.7	35.5	38.9	41.4	21
22	8.64	9.54	10.98	12.34	14.04	17.24	21.3	26.0	30.8	33.9	36.8	40.3	42.8	22
23	9.26	10.20	11.69	13.09	14.85	18.14	22.3	27.1	32.0	35.2	38.1	41.6	44.2	23
24	9.89	10.86	12.40	13.85	15.66	19.04	23.3	28.2	33.2	36.4	39.4	43.0	45.6	24
25	10.52	11.52	13.12	14.61	16.47	19.94	24.3	29.3	34.4	37.7	40.6	44.3	46.9	25
26	11.16	12.20	13.84	15.38	17.29	20.8	25.3	30.4	35.6	38.9	41.9	45.6	48.3	26
27	11.81	12.88	14.57	16.15	18.11	21.7	26.3	31.5	36.7	40.1	43.2	47.0	49.6	27
28	12.46	13.56	15.31	16.93	18.94	22.7	27.3	32.6	37.9	41.3	44.5	48.3	51.0	28
29	13.12	14.26	16.05	17.71	19.77	23.6	28.3	33.7	39.1	42.6	45.7	49.6	52.3	29
30	13.79	14.95	16.79	18.49	20.6	24.5	29.3	34.8	40.3	43.8	47.0	50.9	53.7	30
40	20.7	22.2	24.4	26.5	29.1	33.7	39.3	45.6	51.8	55.8	59.3	63.7	66.8	40
50	28.0	29.7	32.4	34.8	37.7	42.9	49.3	56.3	63.2	67.5	71.4	76.2	79.5	50
60	35.5	37.5	40.5	43.2	46.5	52.3	59.3	67.0	74.4	79.1	83.3	88.4	92.0	60
70	43.3	45.4	48.8	51.7	55.3	61.7	69.3	77.6	85.5	90.5	95.0	100.4	104.2	70
80	51.2	53.5	57.2	60.4	64.3	71.1	79.3	88.1	96.6	101.9	106.6	112.3	116.3	80
90	59.2	61.8	65.6	69.1	73.3	80.6	89.3	98.6	107.6	113.1	118.1	124.1	128.3	90
100	67.3	70.1	74.2	77.9	82.4	90.1	99.3	109.1	118.5	124.3	129.6	135.8	140.2	100
Z_α	−2.58	−2.33	−1.96	−1.64	−1.28	−0.674	0.000	0.674	1.282	1.645	1.960	2.33	2.58	Z_α

For $\phi > 100$ take $\chi^2 = \frac{1}{2}[Z_\alpha + \sqrt{2\phi - 1}]^2$. Z_α is the standardized normal deviate corresponding to the α level of significance, and is shown in the bottom of the table.

SOURCE: This table is abridged from "Table of percentage points of the χ^2 distribution" by Catherine M. Thompson, *Biometrika*, Vol. 32 (1941), pp. 187–191, and is published here by permission of the author and editor of Biometrika.

Degrees of freedom for denominator	Degrees of freedom for numerator 1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	161	200	216	225	230	234	237	239	241	242	244	246	248	249	250	251	252	253	254
2	18.5	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.5	19.5	19.5	19.5	19.5	19.5
3	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

*This table is reproduced from M. Merrington and C. M. Thompson, "Tables of percentage points of the inverted beta (F) distribution," Biometrika 33 (1943), by permission of the Biometrika trustees.

Degrees of freedom for numerator

Degrees of freedom for denominator	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	4,052	5,000	5,403	5,625	5,764	5,859	5,928	5,982	6,023	6,056	6,106	6,157	6,209	6,235	6,261	6,287	6,313	6,339	6,366
2	98.5	99.0	99.2	99.2	99.3	99.3	99.4	99.4	99.4	99.4	99.4	99.4	99.4	99.5	99.5	99.5	99.5	99.5	99.5
3	34.1	30.8	29.5	28.7	28.2	27.9	27.7	27.5	27.3	27.2	27.1	26.9	26.7	26.6	26.5	26.4	26.3	26.2	26.1
4	21.2	18.0	16.7	16.0	15.5	15.2	15.0	14.8	14.7	14.5	14.4	14.2	14.0	13.9	13.8	13.7	13.7	13.6	13.5
5	16.3	13.3	12.1	11.4	11.0	10.7	10.5	10.3	10.2	10.1	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02
6	13.7	10.9	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88
7	12.2	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65
8	11.3	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	4.86
9	10.6	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40	4.31
10	10.0	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69	3.60
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.36
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17
14	8.86	6.51	5.56	5.04	4.70	4.46	4.28	4.14	4.03	3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3.00
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75
17	8.40	6.11	5.19	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66	2.57
19	8.19	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.49
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31	3.17	3.03	2.88	2.80	2.72	2.64	2.55	2.46	2.36
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.07	2.93	2.78	2.70	2.62	2.54	2.45	2.35	2.26
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21
25	7.77	5.57	4.68	4.18	3.86	3.63	3.46	3.32	3.22	3.13	2.99	2.85	2.70	2.62	2.53	2.45	2.36	2.27	2.17
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.60
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53	1.38
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00

*This table is reproduced from M. Merrington and C. M. Thompson, "Tables of percentage points of the inverted beta (F) distribution," Biometrika, Vol. 33 (1943), by permission of the Biometrika trustees.

n	$\binom{n}{0}$	$\binom{n}{1}$	$\binom{n}{2}$	$\binom{n}{3}$	$\binom{n}{4}$	$\binom{n}{5}$	$\binom{n}{6}$	$\binom{n}{7}$	$\binom{n}{8}$	$\binom{n}{9}$	$\binom{n}{10}$
0	1										
1	1	1									
2	1	2	1								
3	1	3	3	1							
4	1	4	6	4	1						
5	1	5	10	10	5	1					
6	1	6	15	20	15	6	1				
7	1	7	21	35	35	21	7	1			
8	1	8	28	56	70	56	28	8	1		
9	1	9	36	84	126	126	84	36	9	1	
10	1	10	45	120	210	252	210	120	45	10	1
11	1	11	55	165	330	462	462	330	165	55	11
12	1	12	66	220	495	792	924	792	495	220	66
13	1	13	78	286	715	1287	1716	1716	1287	715	286
14	1	14	91	364	1001	2002	3003	3432	3003	2002	1001
15	1	15	105	455	1365	3003	5005	6435	6435	5005	3003
16	1	16	120	560	1820	4368	8008	11440	12870	11440	8008
17	1	17	136	680	2380	6188	12376	19448	24310	24310	19448
18	1	18	153	816	3060	8568	18564	31824	43758	48620	43758
19	1	19	171	969	3876	11628	27132	50388	75582	92378	92378
20	1	20	190	1140	4845	15504	38760	77520	125970	167960	184756

If necessary, use the identity $\binom{n}{k} = \binom{n}{n-k}$

n	k	p .05	.10	.15	.20	.25	.30	.35	.40	.45	.50
1	0	.9500	.9000	.8500	.8000	.7500	.7000	.6500	.6000	.5500	.5000
	1	.0500	.1000	.1500	.2000	.2500	.3000	.3500	.4000	.4500	.5000
2	0	.9025	.8100	.7225	.6400	.5625	.4900	.4225	.3600	.3025	.2500
	1	.0950	.1800	.2550	.3200	.3750	.4200	.4550	.4800	.4950	.5000
	2	.0025	.0100	.0225	.0400	.0625	.0900	.1225	.1600	.2025	.2500
3	0	.8574	.7290	.6141	.5120	.4219	.3430	.2746	.2160	.1664	.1250
	1	.1354	.2430	.3251	.3840	.4219	.4410	.4436	.4320	.4084	.3750
	2	.0071	.0270	.0574	.0960	.1406	.1890	.2389	.2880	.3341	.3750
	3	.0001	.0010	.0034	.0080	.0156	.0270	.0429	.0640	.0911	.1250
4	0	.8145	.6561	.5220	.4096	.3164	.2401	.1785	.1296	.0915	.0625
	1	.1715	.2916	.3685	.4096	.4219	.4116	.3845	.3456	.2995	.2500
	2	.0135	.0486	.0975	.1536	.2109	.2646	.3105	.3456	.3675	.3750
	3	.0005	.0036	.0115	.0256	.0469	.0756	.1115	.1536	.2005	.2500
	4	.0000	.0001	.0005	.0016	.0039	.0081	.0150	.0256	.0410	.0625
5	0	.7738	.5905	.4437	.3277	.2373	.1681	.1160	.0778	.0503	.0312
	1	.2036	.3280	.3915	.4096	.3955	.3602	.3124	.2592	.2059	.1562
	2	.0214	.0729	.1382	.2048	.2637	.3087	.3364	.3456	.3369	.3125
	3	.0011	.0081	.0244	.0512	.0879	.1323	.1811	.2304	.2757	.3125
	4	.0000	.0004	.0022	.0064	.0146	.0284	.0488	.0768	.1128	.1562
	5	.0000	.0000	.0001	.0003	.0010	.0024	.0053	.0102	.0185	.0312
6	0	.7351	.5314	.3771	.2621	.1780	.1176	.0754	.0467	.0277	.0156
	1	.2321	.3543	.3993	.3932	.3560	.3025	.2437	.1866	.1359	.0938
	2	.0305	.0984	.1762	.2458	.2966	.3241	.3280	.3110	.2780	.2344
	3	.0021	.0146	.0415	.0819	.1318	.1852	.2355	.2765	.3032	.3125
	4	.0001	.0012	.0055	.0154	.0330	.0595	.0951	.1382	.1861	.2344
	5	.0000	.0001	.0004	.0015	.0044	.0102	.0205	.0369	.0609	.0938
	6	.0000	.0000	.0000	.0001	.0002	.0007	.0018	.0041	.0083	.0156
7	0	.6983	.4783	.3206	.2097	.1335	.0824	.0490	.0280	.0152	.0078
	1	.2573	.3720	.3960	.3670	.3115	.2471	.1848	.1306	.0872	.0547
	2	.0406	.1240	.2097	.2753	.3115	.3177	.2985	.2613	.2140	.1641
	3	.0036	.0230	.0617	.1147	.1730	.2269	.2679	.2903	.2918	.2734
	4	.0002	.0026	.0109	.0287	.0577	.0972	.1442	.1935	.2388	.2734
	5	.0000	.0002	.0012	.0043	.0115	.0250	.0466	.0774	.1172	.1641
	6	.0000	.0000	.0001	.0004	.0013	.0036	.0084	.0172	.0320	.0547
	7	.0000	.0000	.0000	.0000	.0001	.0002	.0006	.0016	.0037	.0078
8	0	.6634	.4305	.2725	.1678	.1001	.0576	.0319	.0168	.0084	.0039
	1	.2793	.3826	.3847	.3355	.2670	.1977	.1373	.0896	.0548	.0312
	2	.0515	.1488	.2376	.2936	.3115	.2965	.2587	.2090	.1569	.1094
	3	.0054	.0331	.0839	.1468	.2076	.2541	.2786	.2787	.2568	.2188
	4	.0004	.0046	.0185	.0459	.0865	.1361	.1875	.2322	.2627	.2734

n	x	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50
8	5	.0000	.0004	.0026	.0092	.0231	.0467	.0808	.1239	.1719	.2188
	6	.0000	.0000	.0002	.0011	.0038	.0100	.0217	.0413	.0703	.1094
	7	.0000	.0000	.0000	.0001	.0004	.0012	.0033	.0079	.0164	.0312
	8	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0007	.0017	.0039
9	0	.6302	.3874	.2316	.1342	.0751	.0404	.0207	.0101	.0046	.0020
	1	.2985	.3874	.3679	.3020	.2253	.1556	.1004	.0605	.0339	.0176
	2	.0629	.1722	.2597	.3020	.3003	.2668	.2162	.1612	.1110	.0703
	3	.0077	.0446	.1069	.1762	.2336	.2668	.2716	.2508	.2119	.1641
	4	.0006	.0074	.0283	.0661	.1168	.1715	.2194	.2508	.2600	.2461
	5	.0000	.0008	.0050	.0165	.0389	.0735	.1181	.1672	.2128	.2461
	6	.0000	.0001	.0006	.0028	.0087	.0210	.0424	.0743	.1160	.1641
	7	.0000	.0000	.0000	.0003	.0012	.0039	.0098	.0212	.0407	.0703
	8	.0000	.0000	.0000	.0000	.0001	.0004	.0013	.0035	.0083	.0176
	9	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0008	.0020
10	0	.5987	.3487	.1969	.1074	.0563	.0282	.0135	.0060	.0025	.0010
	1	.3151	.3874	.3474	.2684	.1877	.1211	.0725	.0403	.0207	.0098
	2	.0746	.1937	.2759	.3020	.2816	.2335	.1757	.1209	.0763	.0439
	3	.0105	.0574	.1298	.2013	.2503	.2668	.2522	.2150	.1665	.1172
	4	.0010	.0112	.0401	.0881	.1460	.2001	.2377	.2508	.2384	.2051
	5	.0001	.0015	.0085	.0264	.0584	.1029	.1536	.2007	.2340	.2461
	6	.0000	.0001	.0012	.0055	.0162	.0368	.0689	.1115	.1596	.2051
	7	.0000	.0000	.0001	.0008	.0031	.0090	.0212	.0425	.0746	.1172
	8	.0000	.0000	.0000	.0001	.0004	.0014	.0043	.0106	.0229	.0439
	9	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0016	.0042	.0098
	10	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0010
11	0	.5688	.3138	.1673	.0859	.0422	.0198	.0088	.0036	.0014	.0005
	1	.3293	.3835	.3248	.2362	.1549	.0932	.0518	.0266	.0125	.0054
	2	.0867	.2131	.2866	.2953	.2581	.1998	.1395	.0887	.0513	.0269
	3	.0137	.0710	.1517	.2215	.2581	.2568	.2254	.1774	.1259	.0806
	4	.0014	.0158	.0536	.1107	.1721	.2201	.2428	.2365	.2060	.1611
	5	.0001	.0025	.0132	.0388	.0803	.1321	.1830	.2207	.2360	.2256
	6	.0000	.0003	.0023	.0097	.0268	.0566	.0985	.1471	.1931	.2256
	7	.0000	.0000	.0003	.0017	.0064	.0173	.0379	.0701	.1128	.1611
	8	.0000	.0000	.0000	.0002	.0011	.0037	.0102	.0234	.0462	.0806
	9	.0000	.0000	.0000	.0000	.0001	.0005	.0018	.0052	.0126	.0269
	10	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0007	.0021	.0054
	11	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0005
12	0	.5404	.2824	.1422	.0687	.0317	.0138	.0057	.0022	.0008	.0002
	1	.3413	.3766	.3012	.2062	.1267	.0712	.0368	.0174	.0075	.0029
	2	.0988	.2301	.2924	.2835	.2323	.1678	.1088	.0639	.0339	.0161
	3	.0173	.0852	.1720	.2362	.2581	.2397	.1954	.1419	.0923	.0537
	4	.0021	.0213	.0683	.1329	.1936	.2311	.2367	.2128	.1700	.1208

n	k	p = .05	.10	.15	.20	.25	.30	.35	.40	.45	.50
12	5	.0002	.0038	.0193	.0532	.1032	.1585	.2039	.2270	.2225	.1934
	6	.0000	.0005	.0040	.0155	.0401	.0792	.1281	.1766	.2124	.2256
	7	.0000	.0000	.0006	.0033	.0115	.0291	.0591	.1009	.1489	.1934
	8	.0000	.0000	.0001	.0005	.0024	.0078	.0199	.0420	.0762	.1208
	9	.0000	.0000	.0000	.0001	.0004	.0015	.0048	.0125	.0277	.0537
	10	.0000	.0000	.0000	.0000	.0000	.0002	.0008	.0025	.0068	.0161
	11	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0010	.0029
	12	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002
13	0	.5133	.2542	.1209	.0550	.0238	.0097	.0037	.0013	.0004	.0001
	1	.3512	.3672	.2774	.1787	.1029	.0540	.0259	.0113	.0045	.0016
	2	.1109	.2448	.2937	.2680	.2059	.1388	.0836	.0453	.0220	.0095
	3	.0214	.0997	.1900	.2457	.2517	.2181	.1651	.1107	.0660	.0349
	4	.0028	.0277	.0838	.1535	.2097	.2337	.2222	.1845	.1350	.0873
	5	.0003	.0055	.0266	.0691	.1258	.1803	.2154	.2214	.1989	.1571
	6	.0000	.0008	.0063	.0230	.0559	.1030	.1546	.1968	.2169	.2095
	7	.0000	.0001	.0011	.0058	.0186	.0442	.0833	.1312	.1775	.2095
	8	.0000	.0000	.0001	.0011	.0047	.0142	.0336	.0656	.1089	.1571
	9	.0000	.0000	.0000	.0001	.0009	.0034	.0101	.0243	.0495	.0873
	10	.0000	.0000	.0000	.0000	.0001	.0006	.0022	.0065	.0162	.0349
	11	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0012	.0036	.0095
	12	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0016
	13	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001
14	0	.4877	.2288	.1028	.0440	.0178	.0068	.0024	.0008	.0002	.0001
	1	.3593	.3559	.2539	.1539	.0832	.0407	.0181	.0073	.0027	.0009
	2	.1229	.2570	.2912	.2501	.1802	.1134	.0634	.0317	.0141	.0056
	3	.0259	.1142	.2056	.2501	.2402	.1943	.1366	.0845	.0462	.0222
	4	.0037	.0349	.0998	.1720	.2202	.2290	.2022	.1549	.1040	.0611
	5	.0004	.0078	.0352	.0860	.1468	.1963	.2178	.2066	.1701	.1222
	6	.0000	.0013	.0093	.0322	.0734	.1262	.1759	.2066	.2088	.1833
	7	.0000	.0002	.0019	.0092	.0280	.0618	.1082	.1574	.1952	.2095
	8	.0000	.0000	.0003	.0020	.0082	.0232	.0510	.0918	.1398	.1833
	9	.0000	.0000	.0000	.0003	.0018	.0066	.0183	.0408	.0762	.1222
	10	.0000	.0000	.0000	.0000	.0003	.0014	.0049	.0136	.0312	.0611
	11	.0000	.0000	.0000	.0000	.0000	.0002	.0010	.0033	.0093	.0222
	12	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0019	.0056
	13	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0009
	14	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001
15	0	.4633	.2059	.0874	.0352	.0134	.0047	.0016	.0005	.0001	.0000
	1	.3658	.3432	.2312	.1319	.0668	.0305	.0126	.0047	.0016	.0005
	2	.1348	.2669	.2856	.2309	.1559	.0916	.0476	.0219	.0090	.0032
	3	.0307	.1285	.2184	.2501	.2252	.1700	.1110	.0634	.0318	.0139
	4	.0049	.0428	.1156	.1876	.2252	.2186	.1792	.1268	.0780	.0417

புள்ளியியல் அட்டவணை - 9உ. (Binomial Probabilities)

n	x	p .05	.10	.15	.20	.25	.30	.35	.40	.45	.50
15	5	.0006	.0105	.0449	.1032	.1651	.2061	.2123	.1859	.1404	.0916
	6	.0000	.0019	.0132	.0430	.0917	.1472	.1906	.2066	.1914	.1527
	7	.0000	.0003	.0030	.0138	.0393	.0811	.1319	.1771	.2013	.1964
	8	.0000	.0000	.0005	.0035	.0131	.0348	.0710	.1181	.1647	.1964
	9	.0000	.0000	.0001	.0007	.0034	.0116	.0298	.0612	.1048	.1527
	10	.0000	.0000	.0000	.0001	.0007	.0030	.0096	.0245	.0515	.0916
	11	.0000	.0000	.0000	.0000	.0001	.0006	.0024	.0074	.0191	.0417
	12	.0000	.0000	.0000	.0000	.0000	.0001	.0004	.0016	.0052	.0139
	13	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0010	.0032
	14	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005
	15	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
16	0	.4401	.1853	.0743	.0281	.0100	.0033	.0010	.0003	.0001	.0000
	1	.3706	.3294	.2097	.1126	.0535	.0228	.0087	.0030	.0009	.0002
	2	.1463	.2745	.2775	.2111	.1336	.0732	.0353	.0150	.0056	.0018
	3	.0359	.1423	.2285	.2463	.2079	.1465	.0888	.0468	.0215	.0085
	4	.0061	.0514	.1311	.2001	.2252	.2040	.1553	.1014	.0572	.0278
	5	.0008	.0137	.0555	.1201	.1802	.2099	.2008	.1623	.1123	.0667
	6	.0001	.0028	.0180	.0550	.1101	.1649	.1982	.1983	.1684	.1222
	7	.0000	.0004	.0045	.0197	.0524	.1010	.1524	.1889	.1969	.1746
	8	.0000	.0001	.0009	.0055	.0197	.0487	.0923	.1417	.1812	.1964
	9	.0000	.0000	.0001	.0012	.0058	.0185	.0442	.0840	.1318	.1746
	10	.0000	.0000	.0000	.0002	.0014	.0056	.0167	.0392	.0755	.1222
	11	.0000	.0000	.0000	.0000	.0002	.0013	.0049	.0142	.0337	.0667
	12	.0000	.0000	.0000	.0000	.0000	.0002	.0011	.0040	.0115	.0278
	13	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0008	.0029	.0085
	14	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0018
	15	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002
	16	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
17	0	.4181	.1668	.0631	.0225	.0075	.0023	.0007	.0002	.0000	.0000
	1	.3741	.3150	.1893	.0957	.0426	.0169	.0060	.0019	.0005	.0001
	2	.1575	.2800	.2673	.1914	.1136	.0581	.0260	.0102	.0035	.0010
	3	.0415	.1556	.2359	.2393	.1893	.1245	.0701	.0341	.0144	.0052
	4	.0076	.0605	.1457	.2093	.2209	.1868	.1320	.0796	.0411	.0182
	5	.0010	.0175	.0668	.1361	.1914	.2081	.1849	.1379	.0875	.0472
	6	.0001	.0039	.0236	.0680	.1276	.1784	.1991	.1839	.1432	.0944
	7	.0000	.0007	.0065	.0267	.0668	.1201	.1685	.1927	.1841	.1484
	8	.0000	.0001	.0014	.0084	.0279	.0644	.1134	.1606	.1883	.1855
	9	.0000	.0000	.0003	.0021	.0093	.0276	.0611	.1070	.1540	.1855
	10	.0000	.0000	.0000	.0004	.0025	.0095	.0263	.0571	.1008	.1484
	11	.0000	.0000	.0000	.0001	.0005	.0026	.0090	.0242	.0525	.0944
	12	.0000	.0000	.0000	.0000	.0001	.0006	.0024	.0081	.0215	.0472
	13	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0021	.0068	.0182
	14	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0004	.0016	.0052

n	k	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50
17	15	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0010
	16	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001
	17	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
18	0	.3972	.1501	.0536	.0180	.0056	.0016	.0004	.0001	.0000	.0000
	1	.3763	.3002	.1704	.0811	.0338	.0126	.0042	.0012	.0003	.0001
	2	.1683	.2835	.2556	.1723	.0958	.0458	.0190	.0069	.0022	.0006
	3	.0473	.1680	.2406	.2297	.1704	.1046	.0547	.0246	.0095	.0031
	4	.0093	.0700	.1592	.2153	.2130	.1681	.1104	.0614	.0291	.0117
	5	.0014	.0218	.0787	.1507	.1988	.2017	.1664	.1146	.0666	.0327
	6	.0002	.0052	.0301	.0816	.1436	.1873	.1941	.1655	.1181	.0708
	7	.0000	.0010	.0091	.0350	.0820	.1376	.1792	.1892	.1657	.1214
	8	.0000	.0002	.0022	.0120	.0376	.0811	.1327	.1734	.1864	.1669
	9	.0000	.0000	.0004	.0033	.0139	.0386	.0794	.1284	.1694	.1855
	10	.0000	.0000	.0001	.0008	.0042	.0149	.0385	.0771	.1248	.1669
	11	.0000	.0000	.0000	.0001	.0010	.0046	.0151	.0374	.0742	.1214
	12	.0000	.0000	.0000	.0000	.0002	.0012	.0047	.0145	.0354	.0708
	13	.0000	.0000	.0000	.0000	.0000	.0002	.0012	.0044	.0134	.0327
	14	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0011	.0039	.0117
	15	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0009	.0031
	16	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0006
	17	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001
	18	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
19	0	.3774	.1351	.0456	.0144	.0042	.0011	.0003	.0001	.0000	.0000
	1	.3774	.2852	.1529	.0685	.0268	.0093	.0029	.0008	.0002	.0000
	2	.1787	.2852	.2428	.1540	.0803	.0358	.0138	.0046	.0013	.0003
	3	.0533	.1796	.2428	.2182	.1517	.0869	.0422	.0175	.0062	.0018
	4	.0112	.0798	.1714	.2182	.2023	.1491	.0909	.0467	.0203	.0074
	5	.0018	.0266	.0907	.1636	.2023	.1916	.1468	.0933	.0497	.0222
	6	.0002	.0069	.0374	.0955	.1574	.1916	.1844	.1451	.0949	.0518
	7	.0000	.0014	.0122	.0443	.0974	.1525	.1844	.1797	.1443	.0961
	8	.0000	.0002	.0032	.0166	.0487	.0981	.1489	.1797	.1771	.1442
	9	.0000	.0000	.0007	.0051	.0198	.0514	.0980	.1464	.1771	.1762
	10	.0000	.0000	.0001	.0013	.0066	.0220	.0528	.0976	.1449	.1762
	11	.0000	.0000	.0000	.0003	.0018	.0077	.0233	.0532	.0970	.1442
	12	.0000	.0000	.0000	.0000	.0004	.0022	.0083	.0237	.0529	.0961
	13	.0000	.0000	.0000	.0000	.0001	.0005	.0024	.0085	.0233	.0518
	14	.0000	.0000	.0000	.0000	.0000	.0001	.0006	.0024	.0082	.0222
	15	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0022	.0074
	16	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0018
	17	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003
	18	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	19	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000

n	k	p: .05	.10	.15	.20	.25	.30	.35	.40	.45	.50
20	0	.3585	.1216	.0388	.0115	.0032	.0008	.0002	.0000	.0000	.0000
	1	.3774	.2702	.1368	.0576	.0211	.0068	.0020	.0005	.0001	.0000
	2	.1887	.2852	.2293	.1369	.0669	.0278	.0100	.0031	.0008	.0002
	3	.0596	.1901	.2428	.2054	.1339	.0716	.0323	.0123	.0040	.0011
	4	.0133	.0898	.1821	.2182	.1897	.1304	.0738	.0350	.0139	.0046
	5	.0022	.0319	.1028	.1746	.2023	.1789	.1272	.0746	.0365	.0148
	6	.0003	.0089	.0454	.1091	.1686	.1916	.1712	.1244	.0746	.0370
	7	.0000	.0020	.0160	.0545	.1124	.1643	.1844	.1659	.1221	.0739
	8	.0000	.0004	.0046	.0222	.0609	.1144	.1614	.1797	.1623	.1201
	9	.0000	.0001	.0011	.0074	.0271	.0654	.1158	.1597	.1771	.1602
	10	.0000	.0000	.0002	.0020	.0099	.0308	.0686	.1171	.1593	.1762
	11	.0000	.0000	.0000	.0005	.0030	.0120	.0336	.0710	.1185	.1602
	12	.0000	.0000	.0000	.0001	.0008	.0039	.0136	.0355	.0727	.1201
	13	.0000	.0000	.0000	.0000	.0002	.0010	.0045	.0146	.0366	.0739
	14	.0000	.0000	.0000	.0000	.0000	.0002	.0012	.0049	.0150	.0370
	15	.0000	.0000	.0000	.0000	.0000	.0000	.0003	.0013	.0049	.0148
	16	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0003	.0013	.0046
	17	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0011
	18	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0002
	19	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	20	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000

Source: This table is extracted from Tables of the Binomial Probability Distribution, National Bureau of Standards, Applied Mathematics Series, 6, U.S. Department of Commerce, 1950.

புள்ளியியல் அட்டவணை - 10அ (Poisson Probabilities)

					λ					
k	.005	.01	.02	.03	.04	.05	.06	.07	.08	.09
0	.9950	.9900	.9802	.9704	.9608	.9512	.9418	.9324	.9231	.9139
1	.0050	.0099	.0192	.0291	.0384	.0476	.0565	.0653	.0738	.0823
2	.0000	.0000	.0002	.0004	.0008	.0012	.0017	.0023	.0030	.0037
3	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001

					λ					
k	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0	.9048	.8187	.7408	.6703	.6065	.5488	.4966	.4493	.4066	.3679
1	.0905	.1637	.2222	.2681	.3033	.3293	.3476	.3595	.3659	.3679
2	.0045	.0164	.0333	.0536	.0758	.0988	.1217	.1438	.1647	.1839
3	.0002	.0011	.0033	.0072	.0126	.0198	.0284	.0383	.0494	.0613
4	.0000	.0001	.0003	.0007	.0016	.0030	.0050	.0077	.0111	.0153
5	.0000	.0000	.0000	.0001	.0002	.0004	.0007	.0012	.0020	.0031
6	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0003	.0005
7	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001

					λ					
k	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0
0	.3329	.3012	.2725	.2466	.2231	.2019	.1827	.1653	.1496	.1353
1	.3662	.3614	.3543	.3452	.3347	.3230	.3106	.2975	.2842	.2707
2	.2014	.2169	.2303	.2417	.2510	.2584	.2640	.2678	.2700	.2707
3	.0738	.0867	.0998	.1128	.1255	.1378	.1496	.1607	.1710	.1804
4	.0203	.0260	.0324	.0395	.0471	.0551	.0636	.0723	.0812	.0902
5	.0045	.0062	.0084	.0111	.0141	.0176	.0216	.0260	.0309	.0361
6	.0008	.0012	.0018	.0026	.0035	.0047	.0061	.0078	.0098	.0120
7	.0001	.0002	.0003	.0005	.0008	.0011	.0015	.0020	.0027	.0034
8	.0000	.0000	.0001	.0001	.0001	.0002	.0003	.0005	.0006	.0009
9	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0002

					λ					
k	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9	3.0
0	.1225	.1108	.1003	.0907	.0821	.0743	.0672	.0608	.0550	.0498
1	.2572	.2438	.2306	.2177	.2052	.1931	.1815	.1703	.1596	.1494
2	.2700	.2681	.2652	.2613	.2565	.2510	.2450	.2384	.2314	.2240
3	.1890	.1966	.2033	.2090	.2138	.2176	.2205	.2225	.2237	.2240
4	.0992	.1082	.1169	.1254	.1336	.1414	.1488	.1557	.1622	.1680
5	.0417	.0476	.0538	.0602	.0668	.0735	.0804	.0872	.0940	.1008
6	.0146	.0174	.0206	.0241	.0278	.0319	.0362	.0407	.0455	.0504
7	.0044	.0055	.0068	.0083	.0099	.0118	.0139	.0163	.0188	.0216
8	.0011	.0015	.0019	.0025	.0031	.0038	.0047	.0057	.0068	.0081
9	.0003	.0004	.0005	.0007	.0009	.0011	.0014	.0018	.0022	.0027
10	.0001	.0001	.0001	.0002	.0002	.0003	.0004	.0005	.0006	.0008
11	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0002	.0002
12	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001

புள்ளியியல் அட்டவணை - 10ஆ (Poisson Probabilities)

k	λ .005	.01	.02	.03	.04	.05	.06	.07	.08	.09
0	.9950	.9900	.9802	.9704	.9608	.9512	.9418	.9324	.9231	.9139
1	.0050	.0099	.0192	.0291	.0384	.0476	.0565	.0653	.0738	.0823
2	.0000	.0000	.0002	.0004	.0008	.0012	.0017	.0023	.0030	.0037
3	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001

k	λ 0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0	.9048	.8187	.7408	.6703	.6065	.5488	.4966	.4493	.4066	.3679
1	.0905	.1637	.2222	.2681	.3033	.3293	.3476	.3595	.3659	.3679
2	.0045	.0164	.0333	.0536	.0758	.0988	.1217	.1438	.1647	.1839
3	.0002	.0011	.0033	.0072	.0126	.0198	.0284	.0383	.0494	.0613
4	.0000	.0001	.0002	.0007	.0016	.0030	.0050	.0077	.0111	.0153
5	.0000	.0000	.0000	.0001	.0002	.0004	.0007	.0012	.0020	.0031
6	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0003	.0005
7	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001

k	λ 1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0
0	.3329	.3012	.2725	.2466	.2231	.2019	.1827	.1653	.1496	.1353
1	.3662	.3614	.3543	.3452	.3347	.3230	.3106	.2975	.2842	.2707
2	.2014	.2169	.2303	.2417	.2510	.2584	.2640	.2678	.2700	.2707
3	.0738	.0867	.0998	.1128	.1255	.1378	.1496	.1607	.1710	.1804
4	.0203	.0260	.0324	.0395	.0471	.0551	.0636	.0723	.0812	.0902
5	.0045	.0062	.0084	.0111	.0141	.0176	.0216	.0260	.0309	.0361
6	.0008	.0012	.0018	.0026	.0035	.0047	.0061	.0078	.0098	.0120
7	.0001	.0002	.0003	.0005	.0008	.0011	.0015	.0020	.0027	.0034
8	.0000	.0000	.0001	.0001	.0001	.0002	.0003	.0005	.0006	.0009
9	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0002

k	λ 2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9	3.0
0	.1225	.1108	.1003	.0907	.0821	.0743	.0672	.0608	.0550	.0498
1	.2572	.2438	.2306	.2177	.2052	.1931	.1815	.1703	.1596	.1494
2	.2700	.2681	.2652	.2613	.2565	.2510	.2450	.2384	.2314	.2240
3	.1890	.1966	.2033	.2090	.2138	.2176	.2205	.2225	.2237	.2240
4	.0992	.1082	.1169	.1254	.1336	.1414	.1488	.1557	.1622	.1680
5	.0417	.0476	.0538	.0602	.0668	.0735	.0804	.0872	.0940	.1008
6	.0146	.0174	.0206	.0241	.0278	.0319	.0362	.0407	.0455	.0504
7	.0044	.0055	.0068	.0083	.0099	.0118	.0139	.0163	.0188	.0216
8	.0011	.0015	.0019	.0025	.0031	.0038	.0047	.0057	.0068	.0081
9	.0003	.0004	.0005	.0007	.0009	.0011	.0014	.0018	.0022	.0027
10	.0001	.0001	.0001	.0002	.0002	.0003	.0004	.0005	.0006	.0008
11	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0002	.0002
12	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001

x	λ 5.1	5.2	5.3	5.4	5.5	5.6	5.7	5.8	5.9	6.0
0	.0061	.0055	.0050	.0045	.0041	.0037	.0033	.0030	.0027	.0025
1	.0311	.0287	.0265	.0244	.0225	.0207	.0191	.0176	.0162	.0149
2	.0793	.0746	.0701	.0659	.0618	.0580	.0544	.0509	.0477	.0446
3	.1348	.1293	.1239	.1185	.1133	.1082	.1033	.0985	.0938	.0892
4	.1719	.1681	.1641	.1600	.1558	.1515	.1472	.1428	.1383	.1339
5	.1753	.1748	.1740	.1728	.1714	.1697	.1678	.1656	.1632	.1606
6	.1490	.1515	.1537	.1555	.1571	.1584	.1594	.1601	.1605	.1606
7	.1086	.1125	.1163	.1200	.1234	.1267	.1298	.1326	.1353	.1377
8	.0692	.0731	.0771	.0810	.0849	.0887	.0925	.0962	.0998	.1033
9	.0392	.0423	.0454	.0486	.0519	.0552	.0586	.0620	.0654	.0688
10	.0200	.0220	.0241	.0262	.0285	.0309	.0334	.0359	.0386	.0413
11	.0093	.0104	.0116	.0129	.0143	.0157	.0173	.0190	.0207	.0225
12	.0039	.0045	.0051	.0058	.0065	.0073	.0082	.0092	.0102	.0113
13	.0015	.0018	.0021	.0024	.0028	.0032	.0036	.0041	.0046	.0052
14	.0006	.0007	.0008	.0009	.0011	.0013	.0015	.0017	.0019	.0022
15	.0002	.0002	.0003	.0003	.0004	.0005	.0006	.0007	.0008	.0009
16	.0001	.0001	.0001	.0001	.0001	.0002	.0002	.0002	.0003	.0003
17	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001	.0001	.0001

x	λ 6.1	6.2	6.3	6.4	6.5	6.6	6.7	6.8	6.9	7.0
0	.0022	.0020	.0018	.0017	.0015	.0014	.0012	.0011	.0010	.0009
1	.0137	.0126	.0116	.0106	.0098	.0090	.0082	.0076	.0070	.0064
2	.0417	.0390	.0364	.0340	.0318	.0296	.0276	.0258	.0240	.0223
3	.0848	.0806	.0765	.0726	.0688	.0652	.0617	.0584	.0552	.0521
4	.1294	.1249	.1205	.1162	.1118	.1076	.1034	.0992	.0952	.0912
5	.1579	.1549	.1519	.1487	.1454	.1420	.1385	.1349	.1314	.1277
6	.1605	.1601	.1595	.1586	.1575	.1562	.1546	.1529	.1511	.1490
7	.1399	.1418	.1435	.1450	.1462	.1472	.1480	.1486	.1489	.1490
8	.1066	.1099	.1130	.1160	.1188	.1215	.1240	.1263	.1284	.1304
9	.0723	.0757	.0791	.0825	.0858	.0891	.0923	.0954	.0985	.1014
10	.0441	.0469	.0498	.0528	.0558	.0588	.0618	.0649	.0679	.0710
11	.0245	.0265	.0285	.0307	.0330	.0353	.0377	.0401	.0426	.0452
12	.0124	.0137	.0150	.0164	.0179	.0194	.0210	.0227	.0245	.0264
13	.0058	.0065	.0073	.0081	.0089	.0098	.0108	.0119	.0130	.0142
14	.0025	.0029	.0033	.0037	.0041	.0046	.0052	.0058	.0064	.0071
15	.0010	.0012	.0014	.0016	.0018	.0020	.0023	.0026	.0029	.0033
16	.0004	.0005	.0005	.0006	.0007	.0008	.0010	.0011	.0013	.0014
17	.0001	.0002	.0002	.0002	.0003	.0003	.0004	.0004	.0005	.0006
18	.0000	.0001	.0001	.0001	.0001	.0001	.0001	.0002	.0002	.0002
19	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001

புள்ளியியல் அட்டவணை - 10ச (Poisson Probabilities)

					λ					
x	7.1	7.2	7.3	7.4	7.5	7.6	7.7	7.8	7.9	8.0
0	.0008	.0007	.0007	.0006	.0006	.0005	.0005	.0004	.0004	.0003
1	.0059	.0054	.0049	.0045	.0041	.0038	.0035	.0032	.0029	.0027
2	.0208	.0194	.0180	.0167	.0156	.0145	.0134	.0125	.0116	.0107
3	.0492	.0464	.0438	.0413	.0389	.0366	.0345	.0324	.0305	.0286
4	.0874	.0836	.0799	.0764	.0729	.0696	.0663	.0632	.0602	.0573
5	.1241	.1204	.1167	.1130	.1094	.1057	.1021	.0986	.0951	.0916
6	.1468	.1445	.1420	.1394	.1367	.1339	.1311	.1282	.1252	.1221
7	.1489	.1486	.1481	.1474	.1465	.1454	.1442	.1428	.1413	.1396
8	.1321	.1337	.1351	.1363	.1373	.1382	.1388	.1392	.1395	.1396
9	.1042	.1070	.1096	.1121	.1144	.1167	.1187	.1207	.1224	.1241
10	.0740	.0770	.0800	.0829	.0858	.0887	.0914	.0941	.0967	.0993
11	.0478	.0504	.0531	.0558	.0585	.0613	.0640	.0667	.0695	.0722
12	.0283	.0303	.0323	.0344	.0366	.0388	.0411	.0434	.0457	.0481
13	.0154	.0168	.0181	.0196	.0211	.0227	.0243	.0260	.0278	.0296
14	.0078	.0086	.0095	.0104	.0113	.0123	.0134	.0145	.0157	.0169
15	.0037	.0041	.0046	.0051	.0057	.0062	.0069	.0075	.0083	.0090
16	.0016	.0019	.0021	.0024	.0026	.0030	.0033	.0037	.0041	.0045
17	.0007	.0008	.0009	.0010	.0012	.0013	.0015	.0017	.0019	.0021
18	.0003	.0003	.0004	.0004	.0005	.0006	.0006	.0007	.0008	.0009
19	.0001	.0001	.0001	.0002	.0002	.0002	.0003	.0003	.0003	.0004
20	.0000	.0000	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0002
21	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001

புள்ளியியல் அட்டவணை - 10உ (Values of e^{-x})

x	e^{-x}	x	e^{-x}	x	e^{-x}	x	e^{-x}
0.0	1.000	2.5	0.082	5.0	0.0067	7.5	0.00055
0.1	0.905	2.6	0.074	5.1	0.0061	7.6	0.00050
0.2	0.819	2.7	0.067	5.2	0.0055	7.7	0.00045
0.3	0.741	2.8	0.061	5.3	0.0050	7.8	0.00041
0.4	0.670	2.9	0.055	5.4	0.0045	7.9	0.00037
0.5	0.607	3.0	0.050	5.5	0.0041	8.0	0.00034
0.6	0.549	3.1	0.045	5.6	0.0037	8.1	0.00030
0.7	0.497	3.2	0.041	5.7	0.0033	8.2	0.00028
0.8	0.449	3.3	0.037	5.8	0.0030	8.3	0.00025
0.9	0.407	3.4	0.033	5.9	0.0027	8.4	0.00023
1.0	0.368	3.5	0.030	6.0	0.0025	8.5	0.00020
1.1	0.333	3.6	0.027	6.1	0.0022	8.6	0.00018
1.2	0.301	3.7	0.025	6.2	0.0020	8.7	0.00017
1.3	0.273	3.8	0.022	6.3	0.0018	8.8	0.00015
1.4	0.247	3.9	0.020	6.4	0.0017	8.9	0.00014
1.5	0.223	4.0	0.018	6.5	0.0015	9.0	0.00012
1.6	0.202	4.1	0.017	6.6	0.0014	9.1	0.00011
1.7	0.183	4.2	0.015	6.7	0.0012	9.2	0.00010
1.8	0.165	4.3	0.014	6.8	0.0011	9.3	0.00009
1.9	0.150	4.4	0.012	6.9	0.0010	9.4	0.00008
2.0	0.135	4.5	0.011	7.0	0.0009	9.5	0.00008
2.1	0.122	4.6	0.010	7.1	0.0008	9.6	0.00007
2.2	0.111	4.7	0.009	7.2	0.0007	9.7	0.00006
2.3	0.100	4.8	0.008	7.3	0.0007	9.8	0.00006
2.4	0.091	4.9	0.007	7.4	0.0006	9.9	0.00005

புள்ளியியல் அட்டவணை - 11அ (The Durbin-Watson d Statistic)

(Significance points of d_l and d_u 5%)

n	k' = 1		k' = 2		k' = 3		k' = 4		k' = 5	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	1.08	1.36	0.95	1.54	0.82	1.75	0.69	1.97	0.56	2.21
16	1.10	1.37	0.98	1.54	0.86	1.73	0.74	1.93	0.62	2.15
17	1.13	1.38	1.02	1.54	0.90	1.71	0.78	1.90	0.67	2.10
18	1.16	1.39	1.05	1.53	0.93	1.69	0.82	1.87	0.71	2.06
19	1.18	1.40	1.08	1.53	0.97	1.68	0.86	1.85	0.75	2.02
20	1.20	1.41	1.10	1.54	1.00	1.68	0.90	1.83	0.79	1.99
21	1.22	1.42	1.13	1.54	1.03	1.67	0.93	1.81	0.83	1.96
22	1.24	1.43	1.15	1.54	1.05	1.66	0.96	1.80	0.86	1.94
23	1.26	1.44	1.17	1.54	1.08	1.66	0.99	1.79	0.90	1.92
24	1.27	1.45	1.19	1.55	1.10	1.66	1.01	1.78	0.93	1.90
25	1.29	1.45	1.21	1.55	1.12	1.66	1.04	1.77	0.95	1.89
26	1.30	1.46	1.22	1.55	1.14	1.65	1.06	1.76	0.98	1.88
27	1.32	1.47	1.24	1.56	1.16	1.65	1.08	1.76	1.01	1.86
28	1.33	1.48	1.26	1.56	1.18	1.65	1.10	1.75	1.03	1.85
29	1.34	1.48	1.27	1.56	1.20	1.65	1.12	1.74	1.05	1.84
30	1.35	1.49	1.28	1.57	1.21	1.65	1.14	1.74	1.07	1.83
31	1.36	1.50	1.30	1.57	1.23	1.65	1.16	1.74	1.09	1.83
32	1.37	1.50	1.31	1.57	1.24	1.65	1.18	1.73	1.11	1.82
33	1.38	1.51	1.32	1.58	1.26	1.65	1.19	1.73	1.13	1.81
34	1.39	1.51	1.33	1.58	1.27	1.65	1.21	1.73	1.15	1.81
35	1.40	1.52	1.34	1.58	1.28	1.65	1.22	1.73	1.16	1.80
36	1.41	1.52	1.35	1.59	1.29	1.65	1.24	1.73	1.18	1.80
37	1.42	1.53	1.36	1.59	1.31	1.66	1.25	1.72	1.19	1.80
38	1.43	1.54	1.37	1.59	1.32	1.66	1.26	1.72	1.21	1.79
39	1.43	1.54	1.38	1.60	1.33	1.66	1.27	1.72	1.22	1.79
40	1.44	1.54	1.39	1.60	1.34	1.66	1.29	1.72	1.23	1.79
45	1.48	1.57	1.43	1.62	1.38	1.67	1.34	1.72	1.29	1.78
50	1.50	1.59	1.46	1.63	1.42	1.67	1.38	1.72	1.34	1.77
55	1.53	1.60	1.49	1.64	1.45	1.68	1.41	1.72	1.38	1.77
60	1.55	1.62	1.51	1.65	1.48	1.69	1.44	1.73	1.41	1.77
65	1.57	1.63	1.54	1.66	1.50	1.70	1.47	1.73	1.44	1.77
70	1.58	1.64	1.55	1.67	1.52	1.70	1.49	1.74	1.46	1.77
75	1.60	1.65	1.57	1.68	1.54	1.71	1.51	1.74	1.49	1.77
80	1.61	1.66	1.59	1.69	1.56	1.72	1.53	1.74	1.51	1.77
85	1.62	1.67	1.60	1.70	1.57	1.72	1.55	1.75	1.52	1.77
90	1.63	1.68	1.61	1.70	1.59	1.73	1.57	1.75	1.54	1.78
95	1.64	1.69	1.62	1.71	1.60	1.73	1.58	1.75	1.56	1.78
100	1.65	1.69	1.63	1.72	1.61	1.74	1.59	1.76	1.57	1.78

புள்ளியியல் அட்டவணை - 11அ (The Durbin-Watson d Statistic)

(Significance points of d_l and d_u 2.5%)

n	k' = 1		k' = 2		k' = 3		k' = 4		k' = 5	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	0.95	1.23	0.83	1.40	0.71	1.61	0.59	1.84	0.48	2.09
16	0.98	1.24	0.86	1.40	0.75	1.59	0.64	1.80	0.53	2.03
17	1.01	1.25	0.90	1.40	0.79	1.58	0.68	1.77	0.57	1.98
18	1.03	1.26	0.93	1.40	0.82	1.56	0.72	1.74	0.62	1.93
19	1.06	1.28	0.96	1.41	0.86	1.55	0.76	1.72	0.66	1.90
20	1.08	1.28	0.99	1.41	0.89	1.55	0.79	1.70	0.70	1.87
21	1.10	1.30	1.01	1.41	0.92	1.54	0.83	1.69	0.73	1.84
22	1.12	1.31	1.04	1.42	0.95	1.54	0.86	1.68	0.77	1.82
23	1.14	1.32	1.06	1.42	0.97	1.54	0.89	1.67	0.80	1.80
24	1.16	1.33	1.08	1.43	1.00	1.54	0.91	1.66	0.83	1.79
25	1.18	1.34	1.10	1.43	1.02	1.54	0.94	1.65	0.86	1.77
26	1.19	1.35	1.12	1.44	1.04	1.54	0.96	1.65	0.88	1.76
27	1.21	1.36	1.13	1.44	1.06	1.54	0.99	1.64	0.91	1.75
28	1.22	1.37	1.15	1.45	1.08	1.54	1.01	1.64	0.93	1.74
29	1.24	1.38	1.17	1.45	1.10	1.54	1.03	1.63	0.96	1.73
30	1.25	1.38	1.18	1.46	1.12	1.54	1.05	1.63	0.98	1.73
31	1.26	1.39	1.20	1.47	1.13	1.55	1.07	1.63	1.00	1.72
32	1.27	1.40	1.21	1.47	1.15	1.55	1.08	1.63	1.02	1.71
33	1.28	1.41	1.22	1.48	1.16	1.55	1.10	1.63	1.04	1.71
34	1.29	1.41	1.24	1.48	1.17	1.55	1.12	1.63	1.06	1.70
35	1.30	1.42	1.25	1.48	1.19	1.55	1.13	1.63	1.07	1.70
36	1.31	1.43	1.26	1.49	1.20	1.56	1.15	1.63	1.09	1.70
37	1.32	1.43	1.27	1.49	1.21	1.56	1.16	1.62	1.10	1.70
38	1.33	1.44	1.28	1.50	1.23	1.56	1.17	1.62	1.12	1.70
39	1.34	1.44	1.29	1.50	1.24	1.56	1.19	1.63	1.13	1.69
40	1.35	1.45	1.30	1.51	1.25	1.57	1.20	1.63	1.15	1.69
45	1.39	1.48	1.34	1.53	1.30	1.58	1.25	1.63	1.21	1.69
50	1.42	1.50	1.38	1.54	1.34	1.59	1.30	1.64	1.26	1.69
55	1.45	1.52	1.41	1.56	1.37	1.60	1.33	1.64	1.30	1.69
60	1.47	1.54	1.44	1.57	1.40	1.61	1.37	1.65	1.33	1.69
65	1.49	1.55	1.46	1.59	1.43	1.62	1.40	1.66	1.36	1.69
70	1.51	1.57	1.48	1.60	1.45	1.63	1.42	1.66	1.39	1.70
75	1.53	1.58	1.50	1.61	1.47	1.64	1.45	1.67	1.42	1.70
80	1.54	1.59	1.52	1.62	1.49	1.65	1.47	1.67	1.44	1.70
85	1.56	1.60	1.53	1.63	1.51	1.65	1.49	1.68	1.46	1.71
90	1.57	1.61	1.55	1.64	1.53	1.66	1.50	1.69	1.48	1.71
95	1.58	1.62	1.56	1.65	1.54	1.67	1.52	1.69	1.50	1.71
100	1.59	1.63	1.57	1.65	1.55	1.67	1.53	1.70	1.51	1.72

புள்ளியியல் அட்டவணை - 12

(Partial Table of Critical Values of U in the Mann-Whitney Test)

Critical Values for One-Tail Test at α = .025 or a Two-Tail Test at α = .05

n_2 \ n_1	9	10	11	12	13	14	15	16	17	18	19	20
1												
2	0	0	0	1	1	1	1	1	2	2	2	2
3	2	3	3	4	4	5	5	6	6	7	7	8
4	4	5	6	7	8	9	10	11	11	12	13	13
5	7	8	9	11	12	13	14	15	17	18	19	20
6	10	11	13	14	16	17	19	21	22	24	25	27
7	12	14	16	18	20	22	24	26	28	30	32	34
8	15	17	19	22	24	26	29	31	34	36	38	41
9	17	20	23	26	28	31	34	37	39	42	45	48
10	20	23	26	29	33	36	39	42	45	48	52	55
11	23	26	30	33	37	40	44	47	51	55	58	62
12	26	29	33	37	41	45	49	53	57	61	65	69
13	28	33	37	41	45	50	54	59	63	67	72	76
14	31	36	40	45	50	55	59	64	67	74	78	83
15	34	39	44	49	54	59	64	70	75	80	85	90
16	37	42	47	53	59	64	70	75	81	86	92	98
17	39	45	51	57	63	67	75	81	87	93	99	105
18	42	48	55	61	67	74	80	86	93	99	106	112
19	45	52	58	65	72	78	85	92	99	106	113	119
20	48	55	62	69	76	83	90	98	105	112	119	127

Critical Values for One-Tail Test at α = .05 or a Two-Tail Test at α = .10

n_2 \ n_1	9	10	11	12	13	14	15	16	17	18	19	20
1											0	0
2	1	1	1	2	2	2	3	3	3	4	4	4
3	3	4	5	5	6	7	7	8	9	9	10	11
4	6	7	8	9	10	11	12	14	15	16	17	18
5	9	11	12	13	15	16	18	19	20	22	23	25
6	12	14	16	17	19	21	23	25	26	28	30	32
7	15	17	19	21	24	26	28	30	33	35	37	39
8	18	20	23	26	28	31	33	36	39	41	44	47
9	21	24	27	30	33	36	39	42	45	48	51	54
10	24	27	31	34	37	41	44	48	51	55	58	62
11	27	31	34	38	42	46	50	54	57	61	65	69
12	30	34	38	42	47	51	55	60	64	68	72	77
13	33	37	42	47	51	56	61	65	70	75	80	84
14	36	41	46	51	56	61	66	71	77	82	87	92
15	39	44	50	55	61	66	72	77	83	88	94	100
16	42	48	54	60	65	71	77	83	89	95	101	107
17	45	51	57	64	70	77	83	89	96	102	109	115
18	48	55	61	68	75	82	88	95	102	109	116	123
19	51	58	65	72	80	87	94	101	109	116	123	130
20	54	62	69	77	84	92	100	107	115	123	130	138

புள்ளியியல் அட்டவணை - 13 (Values of r (Simple Correlation Coefficient) for Different Levels of Significance)

n	.1	.05	.02	.01	.001
1	.98769	.99692	.999507	.999877	.99999
2	.90000	.95000	.98000	.990000	.99900
3	.8054	.8783	.93433	.95873	.99116
4	.7293	.8114	.8822	.91720	.97406
5	.6694	.7545	.8329	.8745	.95074
6	.6215	.7067	.7887	.8343	.92493
7	.5822	.6664	.7498	.7977	.8982
8	.5494	.6319	.7155	.7646	.8721
9	.5214	.6021	.6851	.7348	.8471
10	.4973	.5760	.6581	.7079	.8233
11	.4762	.5529	.6339	.6835	.8010
12	.4575	.5324	.6120	.6614	.7800
13	.4409	.5139	.5923	.6411	.7603
14	.4259	.4973	.5742	.6226	.7420
15	.4124	.4821	.5577	.6055	.7246
16	.4000	.4683	.5425	.5897	.7084
17	.3887	.4555	.5285	.5751	.6932
18	.3783	.4438	.5155	.5614	.6787
19	.3687	.4329	.5034	.5487	.6652
20	.3598	.4227	.4921	.5368	.6524
25	.3233	.3809	.4451	.4869	.5974
30	.2960	.3494	.4093	.4487	.5541
35	.2746	.3246	.3810	.4182	.5189
40	.2573	.3044	.3578	.3932	.4896
45	.2428	.2875	.3384	.3721	.4648
50	.2306	.2732	.3218	.3541	.4433
60	.2108	.2500	.2948	.3248	.4078
70	.1954	.2319	.2737	.3017	.3799
80	.1829	.2172	.2565	.2830	.3568
90	.1726	.2050	.2422	.2673	.3375
100	.1638	.1946	.2301	.2540	.3211

SOURCE: This table is abridged from Table VII of Fisher & Yates: *Statistical Tables for Biological, Agricultural and Medical Research* published by Oliver & Boyd Ltd., Edinburgh, and by permission of the authors and publishers.

புள்ளியியல் அட்டவணை - 14

(Criticial Values of T in the Wilcoxon-Matched-Pairs Test)

N	.025	.01	.005
	Level of Significance for One-Tail Test		
	.05	.02	.01
	Level of Significance for Two-Tail Test		
6	0	—	—
7	2	0	—
8	4	2	0
9	6	3	2
10	8	5	3
11	11	7	5
12	14	10	7
13	17	13	10
14	21	16	13
15	25	20	16
16	30	24	20
17	35	28	23
18	40	33	28
19	46	38	32
20	52	43	38
21	59	49	43
22	66	56	49
23	73	62	55
24	81	69	61
25	89	77	68